

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 15:29:03 ; Search time 162 Seconds
(without alignments)
460.591 Million cell updates/sec

Title: US-10-019-065A-1
Perfect score: 1231
Sequence: 1 QVHGFGQSAWRACSVTCG.....SCGGARQTRGCSDPVPOY 208

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231	100.0	208	5 AAB47770	Aab47770 Human thr
2	1231	100.0	707	6 ABU69134	Abu69134 Human NOV
3	1231	100.0	707	8 ADH72102	Adh72102 Human pro
4	1231	100.0	707	8 ADH08271	Adh08271 Human NOV
5	1231	100.0	1336	5 AAB47771	Aab47771 Human thr
6	1231	100.0	1708	7 ADG39841	Adg39841 Protein s
7	1231	100.0	1708	7 ADG39840	Adg39840 Protein s
8	1231	100.0	1902	6 ABU12084	Abu12084 Human NOV
9	1231	100.0	1902	8 ADH72108	Adh72108 Human pro
10	1231	100.0	5635	5 ABP60991	Abp60991 Novel hum
11	1231	100.0	5636	7 ADJ70089	Adj70089 Human hea
12	1231	100.0	5636	7 ADJ8137	Adj8137 Human hem
13	1231	100.0	5636	8 ADK60205	Adk60205 Angiogene
14	1231	100.0	5636	8 ADK60506	Adk60506 Angiogene
15	1231	100.0	5636	8 ADK60506	Adk60506 Angiogene
16	1230	99.9	1708	7 ADG39844	Adg39844 Protein s
17	1230	99.9	2572	6 ABU12083	Abu12083 Human NOV
18	1230	99.9	2673	8 ADK60225	Adk60225 Angiogene
19	1230	99.9	2673	8 ADK60526	Adk60526 Angiogene
20	1230	99.9	2673	8 ADK60526	Adk60526 Angiogene
21	1230	99.9	3645	7 ADH73149	Adh73149 Angiogene
22	1230	99.9	4495	6 ADH83136	Adh83136 Human hem
23	1230	99.9	4495	8 ABU69135	Abu69135 Human NOV
24	1230	99.9	4495	8 ADH72106	Adh72106 Human pro
25	1221	99.2	712	8 ADH08273	Adh08273 Human NOV
					Adh72104 Human pro

ALIGNMENTS

RESULT 1

AAB47770
ID AAB47770 standard; peptide; 208 AA.

AC AAB47770;
XX
DT 04-MAR-2002 (first entry)
XX
DE Human thrombospondin protein, BTL.012, fragment 654-861.
XX
KW Human; thrombospondin; BTL.012; thrombospondin repeat domain; modulation;
KW angiogenesis; cancer; metastasis; diabetic retinopathy;
KW macular degeneration; cardiovascular disease; wound.
XX
OS Homo sapiens.
XX
PN WO200174852-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US010222.
XX
PR 31-MAR-2000; 2000US-0266300P.
XX
(FARB) BAYER CORP.
XX
PI Chen J, Chen D, Zolotarev A, Davies CJ, Wetzel GD;
PI Dubois-Stringfellow NA;
XX
DR WPI; 2002-066297/09.
XX
PT Novel protein designated BTL.012 has a thrombospondin repeat and inhibits
PT angiogenesis, and is useful to treat cancer, diabetic retinopathy,
PT macular degeneration, cardiovascular disease and wounds.
XX
PS Claim 2; Page 35; 54pp; English.
XX
CC This sequence represents a fragment of the human thrombospondin protein,
CC BTL.012. This peptide, comprising a thrombospondin repeat domain, may be
CC used to modulate angiogenesis at a site. This BTL.012 fragment may be
CC used for diagnosing, preventing or treating a medical condition,
CC particularly cancer, metastasis, diabetic retinopathy, macular
CC degeneration, cardiovascular disease or a wound
XX
SQ Sequence 208 AA;

Query Match 100.0%; Score 1231; DB 5; Length 208;
Best Local Similarity 100.0%; Pred. No. 9.9e-89;

Matches	208;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSLEMRNCQNKPCPYD	60						
Db	1	QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSLEMRNCQNKPCPYD	60						
QY	61	GSWSESLWEECTRCGRGNQTRTRTCNNPVSQHGGRPCGNAVEIIMCNIRPCPVHGAW	120						
Db	61	GSWSESLWEECTRCGRGNQTRTRTCNNPVSQHGGRPCGNAVEIIMCNIRPCPVHGAW	120						
QY	121	SAWQPMGTCSGCGTQTRARLNNPPPAFGGSCDGAETQMVNCNRCPIHGKATW	180						
Db	121	SAWQPMGTCSGCGTQTRARLNNPPPAFGGSCDGAETQMVNCNRCPIHGKATW	180						
QY	181	ASWSACSVSCGGARQTRGSDPVPQY	208						
Db	181	ASWSACSVSCGGARQTRGSDPVPQY	208						
RESULT 2									
ABU69134									
ID	ABU69134 standard; protein; 707 AA.								
XX									
AC	ABU69134;								
XX									
DT	02-JUN-2003 (first entry)								
XX									
DE	Human NOVX polypeptide #9.								
XX									
KW	Human; NOVX; metabolic disorder; diabetes; infectious disease; obesity;								
KW	anorexia; cancer; cardiovascular disorder; asthma; neurogenesis;								
KW	neurodegenerative disorder; epilepsy; immune disorder; osteoarthritis;								
KW	haematopoietic disorder; inflammatory skin disorder; dyslipidemia;								
KW	haematopoiesis; wound healing; angiogenesis; bacterial infection;								
KW	viral infection; fungal infection; helminthic infection; atherosclerosis;								
KW	protozoal infection; hypertension.								
OS	Homo sapiens.								
XX									
PN	WO200290504-A2.								
XX									
PD	14-NOV-2002								
XX									
PF	02-MAY-2002; 2002WO-US014342.								
XX									
PR	03-MAY-2001;	2001US-0288395P.							
PR	04-MAY-2001;	2001US-0288900P.							
PR	07-MAY-2001;	2001US-0289087P.							
PR	14-MAY-2001;	2001US-0290753P.							
PR	15-MAY-2001;	2001US-0291189P.							
PR	16-MAY-2001;	2001US-0291243P.							
PR	18-MAY-2001;	2001US-0292001P.							
PR	21-MAY-2001;	2001US-0292374P.							
PR	22-MAY-2001;	2001US-0292587P.							
PR	23-MAY-2001;	2001US-0293107P.							
PR	29-MAY-2001;	2001US-0294110P.							
PR	30-MAY-2001;	2001US-0294434P.							
PR	31-MAY-2001;	2001US-0294827P.							
PR	18-JUN-2001;	2001US-0298988P.							
PR	31-JUL-2001;	2001US-0308901P.							
PR	17-AUG-2001;	2001US-0313388P.							
PR	21-AUG-2001;	2001US-0313851P.							
PR	21-AUG-2001;	2001US-0313937P.							
PR	17-SEP-2001;	2001US-0322701P.							
PR	17-SEP-2001;	2001US-0322802P.							
PR	25-SEP-2001;	2001US-0324757P.							
PR	27-SEP-2001;	2001US-0325314P.							
PR	27-SEP-2001;	2001US-0325682P.							
PR	21-NOV-2001;	2001US-0332129P.							
PR	03-DEC-2001;	2001US-0336882P.							
PR	14-DEC-2001;	2001US-0340305P.							
PR	01-MAY-2002;	2002US-00138588.							

PA	(CURA-) CURAGEN CORP.
XX	
PI	Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
PI	Chapoval A, Edinger S, Gerlach V, Gorman L, Gunther E, Guo X;
PI	Kekuda R, Lepley DW, Li L, Liu X, Malvankar UM, Miller CE;
PI	Milliet I, Padigaru M, Patturajan M, Pena CEA, Rieger DK, Shenoy SG;
PI	Shinkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;
PI	Zerhusen BD;
XX	
DR	WPI: 2003-103512/09.
DR	N-PSDB; ACA10119.
XX	
PT	New isolated NOVX polypeptides and polynucleotides, useful for
PT	preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT	osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT	asthma, or infections.
XX	
PS	Claim 2; Page 101; 340pp; English.
XX	
CC	The invention relates to human NOVX polypeptides and the polynucleotides
CC	encoding them. The polypeptides, polynucleotides and antibodies that bind
CC	immunosepecifically to the polypeptides are useful in the manufacture of a
CC	medicament for treating a syndrome associated with a human disease,
CC	preferably a NOVX-associated disorder. The sequences are useful for
CC	treating, preventing or diagnosing diseases such as metabolic disorders,
CC	diabetes, obesity, infectious diseases (viral, bacterial, fungal,
CC	helminthic, and protozoal), anorexia, cancer, cardiovascular disorders (e.g.
CC	Alzheimer's disease, atherosclerosis), neurodegenerative disorders (e.g.
CC	osteoarthritis, haematopoietic disorders), epilepsy, immune disorders,
CC	asthma and various dyslipidemias. The nucleic acids and polypeptides may
CC	also be used as targets for the identification of small molecules that
CC	modulate or inhibit e.g. neurogenesis, cell differentiation, cell
CC	proliferation, haematopoiesis, wound healing and angiogenesis, and in the
CC	generation of antibodies that bind immunosepecifically to NOVX substances
CC	for use in therapeutic or diagnostic methods. The nucleic acids are
CC	further used as hybridisation probes, and in chromosome mapping, tissue
CC	typing, preventive medicine and pharmacogenomics. Sequences ABU69126-
CC	ABU69171 represent human NOVX polypeptides of the invention
XX	
SQ	Sequence 707 AA;
Query Match 100.0%; Score 1231; DB 6; Length 707;	
Best Local Similarity 100.0%; Pred. No. 3.3e-88;	
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSLEMRNCQNKPCPYD 60
Db	471 QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSLEMRNCQNKPCPYD 530
QY	61 GSWSESLWEECTRCGRGNQTRTRTCNNPVSQHGGRPCGNAVEIIMCNIRPCPVHGAW 120
Db	531 GSWSESLWEECTRCGRGNQTRTRTCNNPVSQHGGRPCGNAVEIIMCNIRPCPVHGAW 590
QY	121 SAWQPMGTCSGCGTQTRARLNNPPPAFGGSCDGAETQMVNCNRCPIHGKATW 180
Db	591 SAWQPMGTCSGCGTQTRARLNNPPPAFGGSCDGAETQMVNCNRCPIHGKATW 650
QY	181 ASWSACSVSCGGARQTRGSDPVPQY 208
Db	651 ASWSACSVSCGGARQTRGSDPVPQY 678
RESULT 3	
ADH72102	
ID	ADH72102 standard; protein; 707 AA.
XX	
AC	ADH72102;
XX	
DT	25-MAR-2004 (first entry)
XX	
DE	Human protein of the invention NOV43a SEQ ID NO:998.
XX	

QY 181 ASWSACSVSCGGARQRTGCSDFVPOY 208
 Db 651 ASWSACSVSCGGARQRTGCSDFVPOY 678

RESULT 4
 ADO08271
 ID ADO08271 standard; protein; 707 AA.
 AC ADO08271;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human NOVX polypeptide #9.
 XX
 KW Human; NOVX; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
 KW obesity; cancer; diabetes; haemophilia; graft-versus-host disease; AIDS;
 KW asthma; Crohn's disease; multiple sclerosis; infection; anorexia;
 KW cancer-associated cachexia; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; haematopoietic disorder;
 KW wasting disorder.
 XX
 OS Homo sapiens.
 XX
 PN US2004018594-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 01-MAY-2002; 2002US-00138588.
 XX
 PR 03-MAY-2001; 2001US-0288395P.
 PR 04-MAY-2001; 2001US-0288900P.
 PR 07-MAY-2001; 2001US-0289087P.
 PR 14-MAY-2001; 2001US-0290753P.
 PR 15-MAY-2001; 2001US-0291189P.
 PR 16-MAY-2001; 2001US-0291243P.
 PR 18-MAY-2001; 2001US-0292001P.
 PR 21-MAY-2001; 2001US-0292374P.
 PR 22-MAY-2001; 2001US-0292587P.
 PR 23-MAY-2001; 2001US-0293107P.
 PR 29-MAY-2001; 2001US-0294110P.
 PR 30-MAY-2001; 2001US-0294434P.
 PR 31-MAY-2001; 2001US-0294827P.
 PR 31-JUL-2001; 2001US-030901P.
 PR 17-AUG-2001; 2001US-031388P.
 PR 21-AUG-2001; 2001US-0313851P.
 PR 21-AUG-2001; 2001US-0313937P.
 PR 17-SEP-2001; 2001US-0322701P.
 PR 17-SEP-2001; 2001US-0322802P.
 PR 25-SEP-2001; 2001US-0324757P.
 PR 27-SEP-2001; 2001US-0325314P.
 PR 27-SEP-2001; 2001US-0325682P.
 PR 21-NOV-2001; 2001US-0332129P.
 PR 03-DEC-2001; 2001US-0336882P.
 PR 14-DEC-2001; 2001US-0340305P.
 XX
 (ALSO/) ALSBROOK J P.
 PA (ANDE/) ANDERSON D W.
 PA (BOLD/) BOLDOG F L.
 PA (BURG/) BURGESS C E.
 PA (CASM/) CASMAN S J.
 PA (CHAP/) CHAPOVAL A.
 PA (EDIN/) EDINGER S R.
 PA (GERL/) GERLACH V.
 PA (GORM/) GORMAN L.
 PA (GUNT/) GUNTER E.
 PA (GUOX/) GUO X S.
 PA (KEKU/) KEKUDA R.
 PA (LEPL/) LEFLEY D M.
 PA (LILL/) LI L.
 PA (LIUX/) LIU X.
 PA (MALY/) MALYANKAR U M.
 PA (MILL/) MILLER C E.

PA (MILL/) MILLET I.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENNA C E A.
 PA (RIEG/) RIEGER D K.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (SPIT/) SPYTEK K A.
 PA (TAUP/) TAUPIER R J.
 PA (VERN/) VERNET C A M.
 PA (VOSS/) VOSS E Z.
 PA (ZERH/) ZERHUSEN B D.
 XX
 PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
 PI Chapoval A, Edinger SR, Gerlach V, Gorman L, Gunther E, Guo XS;
 PI Kekuda R, Lepley DM, Li L, Liu X, Malvankar UM, Miller CE;
 PI Millet I, Padigar M, Patturajan M, Pena CEA, Rieger DK, Shenoy SG;
 PI Shimkets RA, Spytek KA, Taupier RJ, Vernet CAM, Voss EZ;
 PI Zerhusen BD;
 XX WPI; 2004-122037/12.
 DR N-PSDB; ADO08270.
 XX
 PT New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, diabetes,
 PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing
 PT or pharmacogenomics.
 XX
 PS Claim 2; SEQ ID NO 18; 219pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The polypeptides, polynucleotides and antibodies that bind
 CC immunospecifically to the polypeptides are useful in diagnosing, treating
 CC or preventing NOVX-associated disorders such as cardiomyopathy,
 CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,
 CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,
 CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,
 CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's
 CC disease), haematopoietic disorders and wasting disorders. The
 CC polynucleotides are also used as hybridisation probes, in chromosome
 CC mapping and in tissue typing. The polypeptides are also useful as
 CC vaccines. This sequence represents a human NOVX polypeptide of the
 CC invention.
 XX
 SQ Sequence 707 AA;
 Query Match 100.0%; Score 1231; DB 8; Length 707;
 Best Local Similarity 100.0%; Pred. No. 3.3e-88;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QVHGFSQMSAWRACSVTCGKGIQKESRLCNQPLPANGKPCQGSQDLEMRNQNKPVPYD 60
 Db 471 QVHGFSQMSAWRACSVTCGKGIQKESRLCNQPLPANGKPCQGSQDLEMRNQNKPVPYD 530
 QY 61 GSWSESLWEECTRCGRGNQTRTTCNNPSVQHGRCPCGNVAILMNCNIRPCVPHGAW 120
 Db 531 GSWSESLWEECTRCGRGNQTRTTCNNPSVQHGRCPCGNVAILMNCNIRPCVPHGAW 590
 QY 121 SAWQPMGTCSGCGKTQTRARLNNPPAFGSGYCDGAETQMOVCNERNCPHGWATW 180
 Db 591 SAWQPMGTCSGCGKTQTRARLNNPPAFGSGYCDGAETQMOVCNERNCPHGWATW 650
 QY 181 ASWSACSVSCGGARQRTGCSDFVPOY 208
 Db 651 ASWSACSVSCGGARQRTGCSDFVPOY 678

RESULT 5
 AAB47771
 ID AAB47771 standard; protein; 1336 AA.
 XX
 AC AAB47771;
 XX

DT 04-MAR-2002 (first entry)
XX Human thrombospondin protein, BTL.012.
DE
XX
XX Human; thrombospondin; BTL.012; thrombospondin repeat domain; modulation;
KW angiogenesis; cancer; metastasis; diabetic retinopathy;
KW macular degeneration; cardiovascular disease; wound.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH 657..711
FT Domain /label= thrombospondin domain type I repeat - a
FT Domain 714..768
FT Domain /label= thrombospondin domain type I repeat - b
FT Domain 771..825
FT Domain /label= thrombospondin domain type I repeat - c
FT Domain 828..892
FT Domain /label= thrombospondin domain type I repeat - d
FT Domain 895..939
FT Domain /label= thrombospondin domain type I repeat - e
FT Domain 942..996
FT Domain /label= thrombospondin domain type I repeat - f
FT Misc-difference 1325
FT /label= Encoded by GAA
XX
XX WO200174852-A2.
XX
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US010222.
XX
XX 31-MAR-2000; 2000US-0266300P.
XX
XX (FARB) BAYER CORP.
XX
XX Chen J. Chen D, Zolotorev A, Davies CJ, Wetzel GD;
XX Dubois-Stringfellow NA;
XX
XX WPI; 2002-066297/09.
XX
XX N-PSDB; AAI72024.
XX
XX Novel protein designated BTL.012 has a thrombospondin repeat and inhibits
XX angiogenesis, and is useful to treat cancer, diabetic retinopathy,
XX macular degeneration, cardiovascular disease and wounds.
XX
XX Claim 21; Page 49-53; 54pp; English.
XX
XX This sequence represents a human thrombospondin protein, BTL.012. A
XX fragment of this protein, comprising a thrombospondin repeat domain, may
XX be used to modulate angiogenesis at a site. The BTL.012 fragment may be
XX used for diagnosing, preventing or treating a medical condition,
XX particularly cancer, metastasis, diabetic retinopathy, macular
XX degeneration, cardiovascular disease or a wound
XX
XX Sequence 1336 AA;
XX
XX Query Match 100.0%; Score 1231; DB 5; Length 1336;
XX Best Local Similarity 100.0%; Pred. No. 6.3e-88;
XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 QVHGFGSQAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSDELNRNCKNPKCPVD 60
XX 654 QVHGFGSQAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSDELNRNCKNPKCPVD 713
XX
XX 61 GSWSESLWEECTRSCGNGNTRTCTCNPNPSVQHGRPCGNAVEIIMCNIRPCVPHGAW 120
XX 714 GSWSESLWEECTRSCGNGNTRTCTCNPNPSVQHGRPCGNAVEIIMCNIRPCVPHGAW 773
XX
XX 121 SAWQPMGTCSGCGKGTQTFARLCNNPPPAFGSSYCDGAETQMVCNERNCPFHGKWATW 180
XX 774 SAWQPMGTCSGCGKGTQTFARLCNNPPPAFGSSYCDGAETQMVCNERNCPFHGKWATW 833
XX
XX

QY 181 ASWSACSVCSCGGARQRTGCGSDPVPQY 208
DB 834 ASWSACSVCSCGGARQRTGCGSDPVPQY 861
RESULT 6
ADG39841
ID ADG39841 standard; protein; 1708 AA.
XX
XX AC ADG39841;
XX
XX 26-FEB-2004 (first entry)
XX
XX Protein similar to human NOV9 #2.
XX
XX Human; NOVX; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; pulmonary stenosis; scleroderma; obesity;
KW metabolic disturbance; obesity; transplantation; adrenoleukodystrophy;
KW congenital adrenal hyperplasia; prostate cancer; diabetes;
KW metabolic disorder; neoplasm; adenocarcinoma; fertility; haemophilia;
KW graft versus host disease; AIDS; bronchial asthma; Crohn's disease;
KW multiple sclerosis; infectious disease; anorexia;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW wasting disorder.
XX
XX Homo sapiens.
XX
XX US2003203843-A1.
XX
XX 30-OCT-2003.
XX
XX 11-APR-2002; 2002US-00120801.
XX
XX 20-APR-2001; 2001US-0285609P.
XX 23-APR-2001; 2001US-0285748P.
XX 24-APR-2001; 2001US-0286068P.
XX 25-APR-2001; 2001US-0286292P.
XX 03-MAY-2001; 2001US-0288334P.
XX 16-MAY-2001; 2001US-0291241P.
XX 14-SEP-2001; 2001US-0322284P.
XX
XX (PENA/) PENA C E A.
XX (GUOX/) GUO X.
XX (SHIM/) SHIMKETS R A.
XX (PADI/) PADIGARU M.
XX (KEKU/) KEKUDA R.
XX (SPYT/) SPYTEK K A.
XX (MEHR/) MEHRABAN F.
XX (TOPP/) TOPPER J N.
XX (MALY/) MALYANKAR U M.
XX (WASS/) WASSERMAN S M.
XX (EDIN/) EDINGER S R.
XX (SMIT/) SMITHSON G.
XX (GUNT/) GUNTHER E.
XX (KOMU/) KOMUVES L.
XX
XX Pena CEA, Guo X, Shimkets RA, Padigaru M, Kekuda R, Spytek KA;
XX Mehraban F, Topper JN, Malyankar UM, Wasserman SM, Edinger SR;
XX Smithson G, Gunther E, Komuves L;
XX
XX WPI; 2003-900671/82.
XX
XX New NOVX polypeptides and nucleic acids, useful for diagnosing or
XX treating e.g. cardiomyopathy, atherosclerosis, hypertension, scleroderma,
XX obesity, prostate cancer, AIDS, bronchial asthma, Crohn's disease, or
XX multiple sclerosis.
XX
XX Disclosure; SEQ ID NO 73; 215pp; English.
XX
XX The invention relates to a new isolated polypeptide comprising an amino
XX acid sequence selected from 17 fully defined human NOVX sequences (even
XX SEQ ID NOS between ADG39770 and ADG39802), a mature form of the NOVX
XX

CC amino acid or a variant of NOVX, where one or more amino acid residue in
 CC the variant differs in no more than 15% of the amino acid residue of
 CC NOVX. Also included are an isolated nucleic acid (NA) molecule
 CC (comprising a nucleic acid sequence encoding a NOVX polypeptide above
 CC (odd SEQ ID NOS between ADG39769 and ADG39801), a nucleic acid fragment
 CC encoding at least a portion of a NOVX polypeptide and a complement of
 CC NOVX NA), a vector comprising NOVX NA, a cell comprising the vector, an
 CC antibody that immunospecifically binds to NOVX, a method for determining
 CC the presence or amount of NOVX or NOVX NA in a sample, a method for
 CC identifying an agent that binds to NOVX, a method for identifying an
 CC agent that modulates the expression or activity of NOVX, a method for
 CC modulating the activity of NOVX, a method for treating or preventing a
 CC NOVX-associated disorder, a method for screening for a modulator of
 CC activity or of latency or predisposition to a NOVX-associated disorder, a
 CC method for determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX or NOVX NA in a first mammalian
 CC subject and a method of treating a pathological state in a mammal by
 CC administering NOVX or an antibody that binds to NOVX. The NOVX
 CC polypeptide, nucleic acid or antibody is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease or a
 CC NOVX-associated disorder. The NOVX polypeptides and nucleic acids
 CC encoding them are useful for diagnosing or treating pathologies, diseases
 CC or conditions associated with NOVX sequences, including cardiomyopathy,
 CC atherosclerosis, hypertension, congenital heart defects, pulmonary
 CC stenosis, scleroderma, obesity, metabolic disturbances associated with
 CC hyperplasia, prostate cancer, diabetes, metabolic disorders, neonatal
 CC adenocarcinoma, fertility, haemophilia, graft versus host disease, AIDS,
 CC bronchial asthma, Crohn's disease, multiple sclerosis, infectious disease,
 CC disease, anorexia, neurodegenerative disorders (e.g. Alzheimer's disease,
 CC or Parkinson's disease), immune disorders, haematopoietic disorders,
 CC dyslipidaemias, and wasting disorders associated with chronic diseases.
 CC The polypeptides can be used as immunogens to produce antibodies and as
 CC vaccines. The sequences may further be used in chromosome mapping,
 CC identifying individual from minute biological samples (tissue typing),
 CC and in forensic identification of a biological sample. The present
 CC sequence is a protein showing sequence similarity to a NOVX protein.

XX Sequence 1708 AA;

Query Match 100.0%; Score 1231; DB 7; Length 1708;
 Best Local Similarity 100.0%; Pred. No. 8e-88;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQSAWRACSVTCGKGIQRSLCNQPLPANGKPCQGSDLERNQCQKPCPVD 60
 DB 600 QVHGFSQSAWRACSVTCGKGIQRSLCNQPLPANGKPCQGSDLERNQCQKPCPVD 659

QY 61 GSWSEWSLWEECTRCGNGNQTRTCNNPVSQHGGRPCGNAVEIIMCNIRPCPVHGAW 120
 DB 660 GSWSEWSLWEECTRCGNGNQTRTCNNPVSQHGGRPCGNAVEIIMCNIRPCPVHGAW 719

QY 121 SAWQPGWTCSCGCGTQTRARLCNNPPPAFGSGYCDGAETQMVQCNERNCPHKGWATW 180
 DB 720 SAWQPGWTCSCGCGTQTRARLCNNPPPAFGSGYCDGAETQMVQCNERNCPHKGWATW 779

QY 181 ASWSACSVCSCGGAQRTRGCSDPVQY 208
 DB 780 ASWSACSVCSCGGAQRTRGCSDPVQY 807

RESULT 7

ADG39840

ID ADG39840 standard; protein; 1708 AA.

XX ADG39840;

XX 26-FEB-2004 (first entry)

XX Protein similar to human NOV9 #1.

XX Human; NOVX; cardiomyopathy; atherosclerosis; hypertension;

KW congenital heart defect; pulmonary stenosis; scleroderma; obesity;

KW
 KW
 KW
 KW
 KW
 KW
 KW
 XX

OS Homo sapiens.

XX US2003203843-A1.

PN 30-OCT-2003.

XX 11-APR-2002; 2002US-00120801.

XX 20-APR-2001; 2001US-0285609P.

PR 23-APR-2001; 2001US-0285748P.

PR 24-APR-2001; 2001US-0286068P.

PR 25-APR-2001; 2001US-0286392P.

PR 03-MAY-2001; 2001US-0288344P.

PR 16-MAY-2001; 2001US-0291241P.

PR 14-SEP-2001; 2001US-032284P.

XX (PENA/) PENA C E A.

PA (GUOX/) GUO X.

PA (SHIM/) SHIMKETS R A.

PA (PADI/) PADIGARU M.

PA (KEKU/) KEKUDA R.

PA (SPYT/) SPYTEK K A.

PA (MEHR/) MEHRABAN F.

PA (TOPP/) TOPPER J N.

PA (MALY/) MALYANKAR U M.

PA (WASS/) WASSERMAN S M.

PA (EDIN/) EDINGER S R.

PA (SMIT/) SMITHSON G.

PA (GUNT/) GUNTHER E.

PA (KOMU/) KOMUVES L.

XX Pena CEA, Guo X, Shimkets RA, Padigaru M, Kekuda R, Spytek KA;

PI Mehraban F, Topper JN, Malyankar UM, Wasserman SM, Edinger SR;

PI Smithson G, Gunther E, Komuves L;

XX WPI; 2003-900671/82.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing or

PT treating e.g. cardiomyopathy, atherosclerosis, hypertension, scleroderma,

PT obesity, prostate cancer, AIDS, bronchial asthma, Crohn's disease, or

PT multiple sclerosis.

XX Disclosure; SEQ ID NO 72; 215pp; English.

XX The invention relates to a new isolated polypeptide comprising an amino

CC acid sequence selected from 17 fully defined human NOVX sequences (even

CC SEQ ID NOS between ADG39770 and ADG39802), a mature form of the NOVX

CC amino acid or a variant of NOVX, where one or more amino acid residue in

CC the variant differs in no more than 15% of the amino acid residues of

CC NOVX. Also included are an isolated nucleic acid (NA) molecule

CC (comprising a nucleic acid sequence encoding a NOVX polypeptide above

CC (odd SEQ ID NOS between ADG39769 and ADG39801), a nucleic acid fragment

CC encoding at least a portion of a NOVX polypeptide and a complement of

CC NOVX NA), a vector comprising NOVX NA, a cell comprising the vector, an

CC antibody that immunospecifically binds to NOVX, a method for determining

CC the presence or amount of NOVX or NOVX NA in a sample, a method of

CC identifying an agent that binds to NOVX, a method for identifying an

CC agent that modulates the expression or activity of NOVX, a method for

CC modulating the activity of NOVX, a method for treating or preventing a

CC NOVX-associated disorder, a method for screening for a modulator of

CC activity or of latency or predisposition to a NOVX-associated disorder, a

CC method for determining the presence of or predisposition to a disease

CC associated with altered levels of NOVX or NOVX NA in a first mammalian

CC subject and a method of treating a pathological state in a mammal by

CC

metabolic disturbance; obesity; transplantation; adrenoleukodystrophy;
 congenital adrenal hyperplasia; prostate cancer; diabetes;
 metabolic disorder; neoplasm; adenocarcinoma; fertility; haemophilia;
 graft versus host disease; AIDS; bronchial asthma; Crohn's disease;
 multiple sclerosis; infectious disease; anorexia;
 neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 immune disorder; haematopoietic disorder; dyslipidaemia;
 wasting disorder.

XX

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC administering NOVX or an antibody that binds to NOVX. The NOVX
 CC polypeptide, nucleic acid or antibody is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease or a
 CC NOVX-associated disorder. The NOVX polypeptides and nucleic acids
 CC encoding them are useful for diagnosing or treating pathologies, diseases
 CC or conditions associated with NOVX sequences, including cardiomyopathy,
 CC atherosclerosis, hypertension, congenital heart defects, pulmonary
 CC stenosis, scleroderma, obesity, metabolic disturbances associated with
 CC obesity, transplacental, adrenoleukodystrophy, congenital adrenal
 CC hyperplasia, prostate cancer, diabetes, metabolic disorders, neoplasm,
 CC adenocarcinoma, fertility, haemophilia, graft versus host disease, AIDS,
 CC bronchial asthma, Crohn's disease, multiple sclerosis, infectious
 CC disease, anorexia, neurodegenerative disorders (e.g. Alzheimer's disease,
 CC or Parkinson's disease), immune disorders, haematopoietic disorders,
 CC dyslipidaemias, and wasting disorders associated with chronic diseases.
 CC The polypeptides can be used as immunogens to produce antibodies and as
 CC vaccines. The sequences may further be used in chromosome mapping.
 CC identifying individual from minute biological samples (tissue typing),
 CC and in forensic identification of a biological sample. The present
 CC sequence is a protein showing sequence similarity to a NOVX protein.
 XX
 SQ Sequence 1708 AA;

Query Match 100.0%; Score 1231; DB 7; Length 1708;
 Best Local Similarity 100.0%; Pred. No. 8e-88;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQSGDLEMRNCNKPCPVD 60
 Db 600 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQSGDLEMRNCNKPCPVD 659

QY 61 GSWSEWSLWECTRSCGSGNQTTRTCNNPVSQVHGGRPCGNAVEIINCNRPCPVHGAW 120
 Db 660 GSWSEWSLWECTRSCGSGNQTTRTCNNPVSQVHGGRPCGNAVEIINCNRPCPVHGAW 719

QY 121 SAWQPMWTCSESGKGTOTRRLCNPPPAFGSGYCDGAETQMOVCNERNCPHGWKATW 180
 Db 720 SAWQPMWTCSESGKGTOTRRLCNPPPAFGSGYCDGAETQMOVCNERNCPHGWKATW 779

QY 181 ASWSACSVSCGGARQRTGCSDPVQY 208
 Db 780 ASWSACSVSCGGARQRTGCSDPVQY 807

RESULT 8
 ID ABU12084
 AC ABU12084 standard; protein; 1902 AA.
 XX
 AC ABU12084;
 XX
 DT 19-FEB-2003 (first entry)
 DE Human NOV25c CQ56914-03 protein SEQ ID 98.

XX NOVX; human; antidiabetic; antiarteriosclerotic; anorectic; nootropic;
 KW metabolic; antimicrobial; neuroprotective; antiparkinsonian; cardiant;
 KW antilipemic; cytotatic; immunomodulatory; gene therapy; dyslipidaemia;
 KW cardiomyopathy; metabolic disorder; diabetes; atherosclerosis; obesity;
 KW anorexia; neurodegenerative disorder; Alzheimer's disease; cancer;
 KW Parkinson's disease; haematopoietic disorder; metabolic disturbance;
 KW metabolic syndrome X; wasting disease.
 XX
 OS Homo sapiens.
 XX
 FN W0200281625-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 03-APR-2002; 2002MO-US010366.
 XX
 FR 03-APR-2001; 2001US-0281086P.
 PR 05-APR-2001; 2001US-0281906P.
 PR 06-APR-2001; 2001US-0282020P.

PR 10-APR-2001; 2001US-0282930P.
 PR 12-APR-2001; 2001US-0283444P.
 PR 12-APR-2001; 2001US-0283512P.
 PR 13-APR-2001; 2001US-0283657P.
 PR 13-APR-2001; 2001US-0283678P.
 PR 13-APR-2001; 2001US-0283710P.
 PR 17-APR-2001; 2001US-0284234P.
 PR 19-APR-2001; 2001US-0285325P.
 PR 20-APR-2001; 2001US-0285381P.
 PR 24-APR-2001; 2001US-0286068P.
 PR 25-APR-2001; 2001US-0286292P.
 PR 26-JUN-2001; 2001US-0296692P.
 PR 08-AUG-2001; 2001US-0311003P.
 PR 13-AUG-2001; 2001US-0311973P.
 PR 16-AUG-2001; 2001US-0312901P.
 PR 14-SEP-2001; 2001US-032283P.
 PR 05-OCT-2001; 2001US-0327448P.
 PR 31-DEC-2001; 2001US-0345734P.
 PR 03-JAN-2002; 2002US-0345755P.
 PR 04-FEB-2002; 2002US-0354391P.
 PR 02-APR-2002; 2002US-00114153.
 XX
 XX (CURA-) CURAGEN CORP.

PA Radigaru M, Shenoy SG, Kekuda R, Rastelli L, Mezes PD;
 PI Smithson G, Guo X, Gerlach V, Casman SJ, Boldog FL, Li L;
 PI Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CM, Spyttek KA;
 PI Malyankar UM, Patturajan M, Miller CE, Taupier RJ, Heyes MP, Ju J;
 PI Peyman JA, Catterton E, Macdougall JR, Edinger SR, Stone DJ;
 PI Mazur A;
 XX WPI; 2003-046862/04.
 DR N-PSDB; ABX56304.
 XX

PT New isolated NOVX polypeptide useful for treating cardiomyopathy,
 PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
 PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease and
 PT cancer.
 XX Claim 1; Page 184; 425pp; English.

XX This invention describes novel polypeptides, termed NOVX which have
 CC antidiabetic, antiarteriosclerotic, anorectic, metabolic, antimicrobial,
 CC neuroprotective, antiparkinsonian, antilipemic, cytotatic, nootropic,
 CC cardiant and immunomodulatory activity. The polypeptide and any
 CC antibodies generated from it are useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease
 CC and portions of the polynucleotides encoding NOVX polypeptides are useful
 CC to map the location of NOVX genes on a chromosome, to identify
 CC individuals from minute biological samples, as DNA markers for
 CC restriction fragment length polymorphism (RFLP), and are useful to
 CC prepare polymerase chain reaction primers. The products of the invention
 CC can be used in gene therapy and for treating cardiomyopathy, metabolic
 CC disorders, diabetes, atherosclerosis, obesity, infectious disease,
 CC anorexia, neurodegenerative disorders, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, haematopoietic disorders, and various
 CC dyslipidaemias, metabolic disturbances associated with obesity, metabolic
 CC syndrome X and wasting disorders associated with chronic diseases and
 CC various cancers. ABU12041-ABU12086 represent the polypeptide fragments
 CC encoded by the NOVX polynucleotides represented in ABX56261-ABX56306
 XX
 SQ Sequence 1902 AA;

Query Match 100.0%; Score 1231; DB 6; Length 1902;
 Best Local Similarity 100.0%; Pred. No. 9e-88;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQSGDLEMRNCNKPCPVD 60
 Db 911 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQSGDLEMRNCNKPCPVD 970

CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
 CC further used as hybridisation probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. The present sequence
 CC represents a NOVX polypeptide of the invention.

XX Sequence 1902 AA;

Query Match 100.0%; Score 1231; DB 8; Length 1902;
 Best Local Similarity 100.0%; Pred. No. 9e-88;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QVHGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMCNQKPCPVD 60
 Db 911 QVHGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMCNQKPCPVD 970
 QY 61 GSWSESLWEECTSCGRCNQTRTTCNNPSVQHGRCPCGNAVEIIMCNIIRPCPVHGA 120
 Db 971 GSWSESLWEECTSCGRCNQTRTTCNNPSVQHGRCPCGNAVEIIMCNIIRPCPVHGA 1030
 QY 121 SAMQPMGTCSGCGKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHGWATW 180
 Db 1031 SAMQPMGTCSGCGKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHGWATW 1090
 QY 181 ASWSACSVSCGGARQTRGCDSPVPQY 208
 Db 1091 ASWSACSVSCGGARQTRGCDSPVPQY 1118

RESULT 10
 ABP60991
 ID ABP60991 standard; protein; 5635 AA.

XX AC ABP60991;

DT 10-SEP-2002 (first entry)

XX Novel human protein. SEQ ID 78.

XX Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;
 KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
 KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;
 KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
 KW wound healing disorders; atherosclerosis; Parkinson's disease;
 KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 KW inflammation; neoplastic disease; nervous system disorder;
 KW cardiovascular disorders; pancreatitis; respiratory disorder;
 KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
 KW developmental abnormality; gastrointestinal ulceration; neuropathy;
 KW haematological disease; metabolic disease; sperm dysfunction;
 KW thyroid disorder; hypothyroidism; brain damage; colitis;
 KW cone photo- transduction deficiency; neurological disease; stroke;
 KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
 KW growth abnormality; precocious puberty.

XX Homo sapiens.

XX WO200250105-A1.

XX PD 27-JUN-2002.

XX 17-DEC-2001; 2001WO-US049232.

XX 19-DEC-2000; 2000US-0256710P.

XX 20-DEC-2000; 2000US-0257048P.

XX 09-JAN-2001; 2001US-0260482P.

XX 30-JAN-2001; 2001US-0264922P.

XX 06-FEB-2001; 2001US-0266797P.

XX 19-MAR-2001; 2001US-0276988P.

XX 04-APR-2001; 2001US-0281535P.

XX 08-MAY-2001; 2001US-0283622P.

XX (SWIX) SMITHKLINE BEECHAM CORP.

PA (SWIX) SMITHKLINE BEECHAM PLC.

PA (GLAX) GLAXO GROUP LTD.

XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;

XX Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;

XX WPI; 2002-508784/54.

XX N-PSDB; ABQ86156.

XX Secreted proteins and polynucleotides useful as vaccines for preventing
 XX or treating various diseases e.g. cancer, wounds, atherosclerosis,
 XX Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX Claim 1(a); Page 285-297; 335pp; English.

XX The invention relates to an isolated polypeptide with signal sequences
 XX which allow it to be secreted extracellularly or membrane associated. The
 XX activity of polypeptides of the invention may be described as,
 XX cytosolic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,
 XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 XX cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,
 XX and metabolic. Polypeptides and polynucleotides of the invention are
 XX useful in the treatment, or as a vaccine in the prevention of, cancer,
 XX wound healing disorders, infection, atherosclerosis, Parkinson's disease
 XX and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 XX inflammation, neoplastic diseases, nervous system related disorders and
 XX cardiovascular disorders, pancreatitis, respiratory disorder,
 XX hyperproliferation, systemic autoimmune disease, hyper-immunity,
 XX developmental abnormality, gastrointestinal ulceration, neuropathy,
 XX haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 XX disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
 XX transduction deficiency, neurological diseases, stroke, angiogenesis,
 XX ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 XX trachea, thymus, lymph node and muscular system, obesity, anorexia,
 XX growth abnormalities, and alleviation of precocious puberty. The
 XX sequences given in records ABP60965-ABP61019 represent novel human
 XX proteins of the invention

XX Sequence 5635 AA;

Query Match 100.0%; Score 1231; DB 5; Length 5635;

Best Local Similarity 100.0%; Pred. No. 2.6e-87;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMCNQKPCPVD 60

Db 4527 QVHGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMCNQKPCPVD 4586

QY 61 GSWSESLWEECTSCGRCNQTRTTCNNPSVQHGRCPCGNAVEIIMCNIIRPCPVHGA 120

Db 4587 GSWSESLWEECTSCGRCNQTRTTCNNPSVQHGRCPCGNAVEIIMCNIIRPCPVHGA 4646

QY 121 SAMQPMGTCSGCGKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHGWATW 180

Db 4647 SAMQPMGTCSGCGKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHGWATW 4706

QY 181 ASWSACSVSCGGARQTRGCDSPVPQY 208

Db 4707 ASWSACSVSCGGARQTRGCDSPVPQY 4734

RESULT 11

ADU70089

ID ADU70089 standard; protein; 5636 AA.

XX AC ADU70089;

XX 06-MAY-2004 (first entry)

XX Human heat mitochondrial protein as a therapeutic target SeqID1895.

XX mitochondrial; human; screening assay; diabetes mellitus;

XX Huntington's disease; osteoarthritis;

XX KW

KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
OS Homo sapiens.
XX WO2003087768-A2.
XX PD 23-OCT-2003.
XX PF 04-APR-2003; 2003WO-US010870.
XX PR 12-APR-2002; 2002US-0372843P.
XX PR 17-JUN-2002; 2002US-0389987P.
XX PR 20-SEP-2002; 2002US-0412418P.
XX PA (MITO-) MITOKOR.
XX PA (BUCK-) BUCK INST AGE RES.
XX GHOSH SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX WPI; 2003-845369/78.
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX Claim 1; SEQ ID NO 1895; 180pp; English.
XX This invention relates to novel mitochondrial targets that can be used
XX for therapeutic intervention in treating a disease associated with
XX altered mitochondrial function. Specifically, it refers to a method for
XX identifying proteins of the human heart mitochondrial proteome that are
XX useful for drug screening assays, as well as therapeutic targets. The
XX present invention describes a method for identifying such proteins that
XX can be used in the treatment of various diseases associated with altered
XX mitochondrial function including diabetes mellitus, Huntington's disease,
XX osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX compositions have neuroprotective, nontropic, antidiabetic,
XX anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX cytostatic activities. This polypeptide sequence is a human heart
XX mitochondrial protein of the invention.
SQ Sequence 5636 AA;
Query Match 100.0%; Score 1231; DB 7; Length 5636;
Best Local Similarity 100.0%; Pred. No. 2.6e-87;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QVHGFSONAWRACSVTCGKIQKRLCNQPLPANGKFCQGSLEMRNCQKPCPD 60
DB 4528 QVHGFSONAWRACSVTCGKIQKRLCNQPLPANGKFCQGSLEMRNCQKPCPD 4587
QY 61 GSWSEWSLWECTRCGNGNTRTTCNNPSVQHGRCCEGNVVEIIMCNIRPCVHGAW 120
DB 4598 GSWSEWSLWECTRCGNGNTRTTCNNPSVQHGRCCEGNVVEIIMCNIRPCVHGAW 4647
QY 121 SAWQPGWTCSCGKGTQTRARLNCNPPAPFGSGYCDGAETQMCVNERNCPIHGKWATW 180
DB 4648 SAWQPGWTCSCGKGTQTRARLNCNPPAPFGSGYCDGAETQMCVNERNCPIHGKWATW 4707
QY 181 ASWSACSVSCGGARQRTGCSDPVQY 208
DB 4708 ASWSACSVSCGGARQRTGCSDPVQY 4735
RESULT 12

ADJ83137
ID ADJ83137 standard; protein; 5636 AA.
XX AC ADJ83137;
XX DT 06-MAY-2004 (first entry)
XX DE Human hemocentin protein - SEQ ID 128.
XX KW NOVX; cytostatic; antiarteriosclerotic; antidiabetic; antiasthmatic;
KW antiallergic; antiinflammatory; respiratory; antiarthritic;
KW dermatological; antibacterial; cerebroprotective; vasotropic; cardiant;
KW haemostatic; hypotensive; hepatotropic; neuroprotective; anorectic;
KW nontropic; antitumor; muscular; immunosuppressive; gynaecological;
KW antipsoriatic; endocrine; ophthalmological; osteopathic;
KW antiparkinsonian; anticonvulsant; tranquiliser; analgesic; nephrotropic;
KW antifertility; antilipemic; cardiomyopathy; atherosclerosis; diabetes;
KW cell signal processing; metabolic pathway; asthma; allergy; emphysema;
KW autoimmune; graft-versus-host; arthritis; cancer; stroke; haemophilia;
KW obesity; Alzheimer's; pain; chromosome mapping; tissue typing; human;
KW hemocentin.
XX OS Homo sapiens.
XX PN US2003170630-A1.
XX PD 11-SEP-2003.
XX PF 21-DEC-2001; 2001US-00032189.
XX PR 21-DEC-2000; 2000US-0257495P.
XX PR 22-DEC-2000; 2000US-0258171P.
XX PR 20-FEB-2001; 2001US-0269940P.
XX PR 08-MAR-2001; 2001US-0274192P.
XX PR 22-MAR-2001; 2001US-0277826P.
XX PR 29-MAR-2001; 2001US-0279840P.
XX PR 11-APR-2001; 2001US-0282981P.
XX PR 13-APR-2001; 2001US-0283656P.
XX PR 31-JUL-2001; 2001US-0309247P.
XX PR 10-AUG-2001; 2001US-0311754P.
XX PR 17-AUG-2001; 2001US-0313331P.
XX (ALSO/) ALSOBROOK J P.
XX (TCHE/) TCHERNEV V T.
XX (LIUX/) LIU X.
XX (SPYT/) SPYTEK K A.
XX (ZERH/) ZERHUSEN B D.
XX (PATT/) PATTURAJAN M.
XX (LEPL/) LEPLEY D M.
XX (BURG/) BURGESS C B.
XX (SHIM/) SHIMKETS R A.
XX (GROS/) GROSSE W M.
XX (SZEK/) SZEKERES E S.
XX (VERN/) VERNET C A M.
XX (LILL/) LI L.
XX (CASM/) CASMAN S J.
XX (BOLD/) BOLDOG F L.
XX (GORM/) GORMAN L.
XX (GANG/) GANGOLLI E A.
XX (FERN/) FERNANDES E R.
XX (RIEG/) RIEGER D K.
XX (EDIN/) EDINGER S R.
XX (GUNT/) GUNTHER E.
XX (MILL/) MILLET I.
XX (SCIO/) SCIORE P.
XX (ELLE/) ELLERMAN K.
XX (MACD/) MACDOUGALL J R.
XX (SMIT/) SMITHSON G.
XX Alsobrook JP, Tchernev VT, Liu X, Spytek KA, Zerhusen BD;
XX Patturajan M, Lepley DM, Burgess CE, Shimkets RA, Grosse WM;
XX Szekeres ES, Vernet CAM, Li L, Casman SJ, Boldog FL, Gorman L;
XX Gangolli EA, Fernandes ER, Rieger DK, Edinger SR, Gunther E;

PI Millet I, Sciore P, Ellerman K, Macdougall JR, Smithson G;
 DR WPI; 2003-898249/82.
 XX
 PT New NOVX polypeptides and nucleic acid molecules, useful for diagnosing,
 PT preventing or treating NOVX-associated polypeptide disorder, e.g.
 PT cardiomyopathy, atherosclerosis, diabetes, cancer, Parkinson's disease or
 PT asthma.
 XX
 PS Disclosure; SEQ ID NO 128; 263pp; English.

XX The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide demonstrates cytostatic, antiarteriosclerotic, antidiabetic,
 CC antiasthmatic, anti-allergic, anti-inflammatory, respiratory,
 CC antiarthritic, dermatological, antibacterial, cerebroprotective,
 CC vasotropic, cardiant, haemostatic, hypotensive, hepatotropic,
 CC neuroprotective, anorectic, nootropic, antitumor, muscular,
 CC immunosuppressive, gynaecological, antiparkinsonian, endocrine,
 CC ophthalmological, osteopathic, antiparkinsonian, anticonvulsant,
 CC tranquiliser, analgesic, nephrotropic, antifertility and antilipaeamic
 CC activities. The NOVX polypeptide, nucleic acid or antibody of the
 CC invention may be useful for treating or preventing a NOVX-associated
 CC disorder, such as cardiomyopathy, atherosclerosis, diabetes or a disorder
 CC related to cell signal processing and metabolic pathway modulation.
 CC Furthermore, the NOVX polypeptides may be useful for diagnosing, treating
 CC or preventing diseases such as asthma, allergies, emphysema, autoimmune
 CC disease, graft-versus-host disease, arthritis, cancer, stroke,
 CC haemophilia, obesity, Alzheimer's disease and pain. The nucleic acids may
 CC be used as hybridisation probes, in chromosome mapping, tissue typing,
 CC preventive medicine or pharmacogenomics. The current sequence is that of
 CC a protein of the invention which is related to human NOVX protein.

XX Sequence 5636 AA;
 SQ Query Match 100.0%; Score 1231; DB 7; Length 5636;
 Best Local Similarity 100.0%; Pred. No. 2.6e-87;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVHGFFSQWAWACSVTCGKIGKESRLCNQPLPANGKPCQGSLEMRNCQKPCPD 60
 Db 4528 QVHGFFSQWAWACSVTCGKIGKESRLCNQPLPANGKPCQGSLEMRNCQKPCPD 4587
 Qy 61 GSWSEWSLWEECTSCGRGNQTRTRTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGW 120
 Db 4588 GSWSEWSLWEECTSCGRGNQTRTRTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGW 4647
 Qy 121 SAWQPMWTCSESCGKGTQTRARLCNNPPAFGGSYCDGAETQVQVCNERNCPIHGKWATW 180
 Db 4648 SAWQPMWTCSESCGKGTQTRARLCNNPPAFGGSYCDGAETQVQVCNERNCPIHGKWATW 4707
 Qy 181 ASWSACSVCSCGGARQTRGCDPVPQY 208
 Db 4708 ASWSACSVCSCGGARQTRGCDPVPQY 4735

RESULT 13
 ID ADK60205 standard; protein; 5636 AA.
 XX AC ADK60205;
 XX DT 06-MAY-2004 (first entry)
 XX DE Angiogenesis differentially expressed protein GS-P29.
 KW vasotropic; antirheumatic; antiarthritic; hypotensive; antianginal;
 KW antiinflammatory; cardiant; angiogenesis inhibitor; gene therapy;
 KW angiogenesis; endothelial cell; diagnosis; tumor vascularization;
 KW retinopathy; rheumatoid arthritis; Crohn's disease; atherosclerosis;
 KW ovary hyperstimulation; psoriasis; endometriosis; restenosis;
 KW angioplasty; cicatrization; peripheral vascular disease; hypertension;
 KW vascular inflammation; Raynaud's disease; aneurism; thrombophlebitis;
 KW ischemia; angina; myocardial infarction; chronic heart disease;

KW cardiac congestion; macular degeneration; osteoporosis.
 XX OS Homo sapiens.
 XX PN FR2836687-A1.
 XX PD 05-SEP-2003.
 XX 11-APR-2002; 2002FR-00004546.
 XX 04-MAR-2002; 2002FR-00002717.
 XX (GENE-) GENE SIGNAL.
 XX (ALMA-) AL MAHMOOD S.
 XX Colin S, Schneider C, Al Mahmood S;
 WPI; 2004-013912/02.
 DR N-PSDB; ADK60455.
 XX Compositions for diagnosing, prognosing and treating angiogenic disorders
 PT including tumor vascularization and heart disease, comprise nucleic acid
 PT or polypeptide differentially expressed in angiogenesis.
 XX Claim 7; SEQ ID NO 81; 424pp; French.

XX The invention relates to a novel pharmaceutical composition active on
 CC angiogenesis comprising an endothelial cell nucleic acid whose expression
 CC is induced by an angiogenic factor and inhibited by an angiostatic agent
 CC or its complement or fragment, a polypeptide sequence encoded by the
 CC nucleic acid or its fragment, a molecule capable of inhibiting expression
 CC of the nucleic acid or a molecule which binds to the polypeptide
 CC sequence. The invention is used to diagnose, prognose or treat an
 CC angiogenic disorder in a mammal, particularly a human. The disorder is
 CC particularly tumor vascularization, a retinopathy, rheumatoid arthritis,
 CC Crohn's disease, atherosclerosis, ovary hyperstimulation, psoriasis,
 CC endometriosis associated with neovascularization, restenosis due to
 CC angioplasty, overproduction of tissue due to cicatrization, a peripheral
 CC vascular disease, hypertension, vascular inflammation, Raynaud disease,
 CC aneurism, arterial restenosis, thrombophlebitis, ischemia, angina,
 CC myocardial infarction, chronic heart disease, cardiac congestion or
 CC macular degeneration due to age or osteoporosis. This sequence
 CC corresponds to a protein encoded by a differentially expressed DNA used
 CC in the composition of the invention.

XX Sequence 5636 AA;
 Query Match 100.0%; Score 1231; DB 8; Length 5636;
 Best Local Similarity 100.0%; Pred. No. 2.6e-87;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 QVHGFFSQWAWACSVTCGKIGKESRLCNQPLPANGKPCQGSLEMRNCQKPCPD 60
 Db 4528 QVHGFFSQWAWACSVTCGKIGKESRLCNQPLPANGKPCQGSLEMRNCQKPCPD 4587
 Qy 61 GSWSEWSLWEECTSCGRGNQTRTRTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGW 120
 Db 4588 GSWSEWSLWEECTSCGRGNQTRTRTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGW 4647
 Qy 121 SAWQPMWTCSESCGKGTQTRARLCNNPPAFGGSYCDGAETQVQVCNERNCPIHGKWATW 180
 Db 4648 SAWQPMWTCSESCGKGTQTRARLCNNPPAFGGSYCDGAETQVQVCNERNCPIHGKWATW 4707
 Qy 181 ASWSACSVCSCGGARQTRGCDPVPQY 208
 Db 4708 ASWSACSVCSCGGARQTRGCDPVPQY 4735

RESULT 14
 ID ADK60506 standard; protein; 5636 AA.
 XX AC ADK60506;

XX DT 06-MAY-2004 (first entry)

XX DE Angiogenesis differentially expressed protein GS-P29.

XX KW vasotrophic; anti-rheumatic; antiarthritic; hypotensive; anti-angiogenic; anti-inflammatory; cardiac; angiogenesis inhibitor; gene therapy; angiogenesis; endothelial cell; diagnosis; tumor vascularization; retinopathy; rheumatoid arthritis; Crohn's disease; atherosclerosis; ovary hyperstimulation; psoriasis; endometriosis; restenosis; angioplasty; cicatrization; peripheral vascular disease; hypertension; vascular inflammation; Raynaud's disease; aneurism; thrombophlebitis; ischemia; angina; myocardial infarction; chronic heart disease; cardiac congestion; macular degeneration; osteoporosis.

XX OS Homo sapiens.

XX FN FR2836686-A1.

XX PD 05-SEP-2004.

XX PF 04-MAR-2002; 2002FR-00002717.

XX PR 04-MAR-2002; 2002FR-00002717.

XX PA (GENE-) GENE SIGNAL.

XX PA (ALMA/) AL MAHMOOD S.

XX PI Colin S, Schneider C, Al Mahmood S;

XX DR WPI; 2004-013911/02.

XX DR N-PSDB; ADK60756.

XX PT Compositions containing nucleic acid or polypeptide differentially expressed in angiogenesis are useful to diagnose, prognose and treat angiogenic disorders including tumor vascularization and heart disease.

XX PS Claim 7; SEQ ID NO 81; 405pp; French.

XX CC The invention relates to a novel pharmaceutical composition active on angiogenesis comprising an endothelial cell nucleic acid whose expression is induced by an angiogenic factor and inhibited by an angiostatic agent or its complement or fragment, a polypeptide sequence encoded by the nucleic acid or its fragment, a molecule capable of inhibiting expression of the nucleic acid or a molecule which binds to the polypeptide sequence. The invention is used to diagnose, prognose or treat an angiogenic disorder in a mammal, particularly a human. The disorder is particularly tumor vascularization, a retinopathy, rheumatoid arthritis, Crohn's disease, atherosclerosis, ovary hyperstimulation, psoriasis, endometriosis associated with neovascularization, restenosis due to angioplasty, overproduction of tissue due to cicatrization, a peripheral vascular disease, hypertension, vascular inflammation, Raynaud disease, aneurism, arterial restenosis, thrombophlebitis, ischemia, angina, myocardial infarction, chronic heart disease, cardiac congestion or macular degeneration due to age or osteoporosis. This sequence corresponds to a protein encoded by a differentially expressed DNA used in the composition of the invention.

XX SQ Sequence 5636 AA;

Query Match 100.0%; Score 1231; DB 8; Length 5636;
Best Local Similarity 100.0%; Pred. No. 2.6e-87;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQSAWRACSVTCGKIQKRSRLCNQPLPANGKPKCGSDLEMRNCKNPKCPVD 60
|||||

DB 4528 QVHGFSQSAWRACSVTCGKIQKRSRLCNQPLPANGKPKCGSDLEMRNCKNPKCPVD 4587
|||||

QY 61 GSWSEWSLWEECTRSCGNGNTRTRTCNNPSVQHGRRPCEGNAVEIIMCNIRPCPVHGAW 120
|||||

DB 4588 GSWSEWSLWEECTRSCGNGNTRTRTCNNPSVQHGRRPCEGNAVEIIMCNIRPCPVHGAW 4647
|||||

QY 121 SAWQPMWGTCSBSCGKGTQTRARLCNNPPAFGGVCDGAETQMVQCNERNCPHKGWATW 180
|||||

DB 4648 SAWQPMWGTCSBSCGKGTQTRARLCNNPPAFGGVCDGAETQMVQCNERNCPHKGWATW 4707
|||||

QY 181 ASWSACSVCSCGGARQRTGCSDFVPQY 208
|||||

DB 4708 ASWSACSVCSCGGARQRTGCSDFVPQY 4735
|||||

RESULT 15
ADP73129
ID ADP73129 standard; protein; 5636 AA.
XX AC ADP73129;
XX DT 12-AUG-2004 (first entry)
XX DE Angiogenesis inhibitor human protein sequence, GS-P29.
XX KW Inhibitor; angiogenesis; antisense nucleic acid; immunisation; angiogenic disorder; antiangiogenic; angiogenesis stimulator; cytostatic; dermatological; antiarthritic; anti-rheumatic; anti-inflammatory; vasotrophic; hypotensive; ophthalmological; antipsoriatic; cardiac; gene therapy; antisense gene therapy; tumor vascularisation; retinopathies; rheumatoid arthritis; Crohn's disease; atherosclerosis; ovarian hyperstimulation; psoriasis; endometriosis; restenosis; tissue granulation; peripheral vascular disorder; hypertension; vascular inflammation; Raynaud's disease; aneurism; arterial restenosis; thrombophlebitis; lymphadenopathy; lymphedema; ischaemia; angina; myocardial infarction; chronic heart disease; congestive heart disease; macular degeneration; human.

XX OS Homo sapiens.

XX FN FR2843753-A1.

XX PD 27-FEB-2004.

XX PF 20-JUN-2003; 2003FR-00007507.

XX PR 04-MAR-2002; 2002FR-00002717.

XX PA (GENE/) GENE S.

XX PA (ALMS/) AL M S.

XX PI Colin S, Schneider C, Al MS;

XX DR WPI; 2004-216677/21.

XX DR N-PSDB; ADP73078.

XX PT Antisense nucleic molecule useful as inhibitor of angiogenesis in the treatment of angiogenic disorders, e.g., rheumatoid arthritis, atherosclerosis and endometriosis.

XX PS Claim 1; SEQ ID NO 81; 404pp; French.

XX CC The invention relates to a novel inhibitor of angiogenesis comprising an active substance chosen from at least one of a nucleic acid molecule, an antisense nucleic acid molecule, a polypeptide or an antibody. The invention further comprises: an antisense nucleic acid sequence chosen from any of the sequences provided in the specification; preparation of the antibody comprising in vivo or in vitro immunisation of an immunocompetent animal cell, preferably of a vertebrate and most preferably of a mammal, with at least one of the polypeptide sequences chosen from a sequence provided in the specification; a mammalian expression vector comprising at least one antisense sequence chosen from an antisense nucleic acid provided in the specification; preparation of a genetically modified cell, that over- or under-expresses a gene implicated in an angiogenic disorder, comprising inserting the vector from above into a mammalian cell; a genetically modified cell that over-expresses or under-expresses at least one gene involved in angiogenesis by a nucleic acid sequence chosen from any of ADP73049 to ADP73138, as in the antisense nucleic acid selection of above, or a fragment of any of these; preparation of a cell line that stably expresses an expression

CC vector. The angiogenesis inhibitor has the following activities:
 CC antiangiogenic, angiogenesis stimulator, angiogenesis inhibitor,
 CC cytostatic, dermatological, antiarthritic, antirheumatic,
 CC antiinflammatory, vasotropic, hypotensive, ophthalmological,
 CC antipsoriatic, and cardiant. The novel nucleic acid molecules of the
 CC invention may be used to treat disorders in gene therapy and antisense
 CC gene therapy. The nucleic acid sequences, proteins and antibodies as part
 CC of the therapeutic compositions are useful in treating a disorder of
 CC angiogenesis chosen from: tumour vascularisation, retinopathies,
 CC rheumatoid arthritis, Crohn's disease, atherosclerosis, ovarian
 CC hyperstimulation, psoriasis, endometriosis associated with
 CC neovascularisation, restenosis due to balloon angioplasty, tissue
 CC granulation due to scarification, peripheral vascular disorders,
 CC hypertension, vascular inflammation, Raynaud's disease, aneurism,
 CC arterial restenosis, thrombophlebitis, lymphadenopathy, lymphedema,
 CC ischaemia, angina, myocardial infarction, chronic heart disease,
 CC congestive heart disease, macular degeneration linked to age and
 CC osteoporosis. This sequence represents a human protein for the creation
 CC of an angiogenesis inhibitor of the invention.
 XX
 SQ Sequence 5636 AA;

Query Match 100.0%; Score 1231; DB 8; Length 5636;
 Best Local Similarity 100.0%; Pred. No. 2.6e-87;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 QVHGGSQWAWRACSVTCGKGIQKSRRLCNQPLPANGGKPCQGSDEMRNQNKPVPD 60
 Db 4528 QVHGGSQWAWRACSVTCGKGIQKSRRLCNQPLPANGGKPCQGSDEMRNQNKPVPD 4587
 Qy 61 GSWSESLWEECTSCRGNGTTRTCNNPVSVOHGRPCGEGNAVELIMCNIRPCPVHGAW 120
 Db 4588 GSWSESLWEECTSCRGNGTTRTCNNPVSVOHGRPCGEGNAVELIMCNIRPCPVHGAW 4647
 Qy 121 SAMQPWGTCSESCGKGTQTRARLCNNPPPAFGGSYCDGAETQMVQCNERNCPHGWATW 180
 Db 4648 SAMQPWGTCSESCGKGTQTRARLCNNPPPAFGGSYCDGAETQMVQCNERNCPHGWATW 4707
 Qy 181 ASWSACSVSCGGARQTRCGSDPVPQY 208
 Db 4708 ASWSACSVSCGGARQTRCGSDPVPQY 4735

Search completed: November 17, 2004, 15:52:22
 Job time : 175 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 15:49:06 ; Search time 39 Seconds
(without alignments)
353.696 Million cell updates/sec

Title: US-10-019-065A-1
Perfect score: 1231
Sequence: 1 QVHGGSFQSAWRACSVTCG.....SCGGARQRTGCSDFVPOY 208

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	511.5	41.6	939	4	US-09-854-845-16
2	511.5	41.6	934	4	US-09-854-845-14
3	511.5	41.6	1034	4	US-09-854-845-6
4	511.5	41.6	1049	4	US-09-854-845-2
5	511.5	41.6	1078	4	US-09-854-845-8
6	511.5	41.6	1093	4	US-09-854-845-4
7	511.5	41.6	1136	4	US-09-854-845-12
8	511.5	41.6	1151	4	US-09-854-845-10
9	495.5	40.3	441	3	US-08-985-526-3
10	488.5	39.7	1170	1	US-08-313-288B-20
11	488.5	39.7	1170	4	US-09-657-472-2
12	478.5	38.9	239	5	PCT-US93-01652-1
13	453	36.8	469	1	US-08-313-288B-15
14	453	36.8	1172	1	US-08-313-288B-19
15	417.5	33.9	788	2	US-08-918-914-4
16	398	32.3	479	4	US-09-270-767-46823
17	360	29.2	218	3	US-08-985-526-1
18	288.5	23.4	807	4	US-09-132-769-1
19	288.5	23.4	807	4	US-09-640-173-186
20	288.5	23.4	807	4	US-07-713-550-186
21	288.5	23.4	807	4	US-09-825-294-186
22	288.5	23.4	807	4	US-09-970-966-186
23	287.5	23.4	787	4	US-09-825-294-207
24	287.5	23.4	787	4	US-09-970-966-207
25	287.5	23.4	807	1	US-07-862-021B-10
26	287.5	23.4	807	1	US-08-313-288B-10
27	287.5	23.4	807	4	US-09-132-769-3

28	287.5	23.4	807	4	US-09-132-769-5	Sequence 5, Appli
29	287.5	23.4	807	5	PCT-US93-03164-10	Sequence 10, Appl
30	283.5	23.0	802	1	US-07-862-021B-12	Sequence 12, Appl
31	283.5	23.0	802	1	US-08-313-288B-12	Sequence 12, Appl
32	283.5	23.0	802	5	PCT-US93-03164-12	Sequence 12, Appl
33	281.5	22.9	568	1	US-07-862-021B-14	Sequence 14, Appl
34	281.5	22.9	568	5	PCT-US93-03164-14	Sequence 14, Appl
35	278.5	22.6	898	2	US-08-808-382-5	Sequence 5, Appli
36	278.5	22.6	898	3	US-09-306-902A-5	Sequence 5, Appli
37	276	22.4	2150	4	US-09-321-987B-2	Sequence 2, Appli
38	276	22.4	2165	4	US-09-800-729-155	Sequence 155, App
39	273	22.2	943	2	US-08-808-982-7	Sequence 7, Appli
40	273	22.2	943	3	US-09-306-902A-7	Sequence 7, Appli
41	271	22.0	945	4	US-10-140-002-145	Sequence 146, App
42	268	21.8	1882	3	US-09-369-364A-13	Sequence 13, Appl
43	256.5	20.8	1224	4	US-09-930-872-4	Sequence 4, Appli
44	256.5	20.8	1224	4	US-10-217-774-4	Sequence 4, Appli
45	244.5	19.9	1745	4	US-09-800-729-89	Sequence 89, Appl

ALIGNMENTS

RESULT 1

US-09-854-845-16
; Sequence 16, Application US/09854845
; Patent No. 6750054
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6750054el Human Semaphorin Homologs and Polynucleotides Encod
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 939
; TYPE: PPT
; ORGANISM: homo sapiens
; US-09-854-845-16

Query Match 41.6%; Score 511.5; DB 4; Length 939;
Best Local Similarity 33.0%; Pred. No. 5.7e-35;
Matches 94; Conservative 22; Mismatches 82; Indels 87; Gaps 3;

QY	3	HGFSQSAWFAACSVTCGKIQKSRILCNQPLPANGKPCQGSQSDLEMRNC-QNKPCPVDG	61
Db	467	NGATPSSWALCSTGCGIFQVRQSCSNFAPRHGGRICVKGSRERFCNENPCPVI	526
QY	62	SWSEWSLWEETRSCGRNQTRTRTCNNPSVQHGRCPCGNAVEIIMCNIRPCP	115
Db	527	FWASGWSKSCSSNCGGQMRRACEN-----GNSCLGCGVEFKTCNPEGCPVRNT	580
QY	116	-----	115
Db	591	PWTPLPWNVTGGARQERFTCTRAPLADPHLGFGRRTTRTCPCADGSGSCDTDL	640
QY	116	-----VHGASAWQFWGTCSGSGKGTQTRARLCNNPPFAFGGSCYDGAET	161
Db	641	VEDLLRSGTSPHTVSGWAAWGPWSGSCRDCELGFRVAKRTCTNPEPRNGGLPCVGDAA	700
QY	162	QMVCNERNCFIHKWATWASWACSVSCGGARQRTGCSDFVP	206
Db	701	EYQCNFQACPVRGAMSCWTSWSPCSASCGGHYQRTSCTSPAP	745

DB 838 VEDLLRSGTSPHTVSGGAAWGPWSSCSRDCELGFRVRKKTCTNPEFRNGGLPCVGDAA 897
QY 162 QMVCNERNCPHGWKATWASWSACSVCSSGASQRTGCGSDPVP 206
DB 898 EYQDCNPQACVRGAWSCWTSWSPCSASCGGHHYQRTSCTSPAP 942

RESULT 8
US-09-854-845-10
; Sequence 10, Application US/09854845
; Patent No. 6750054
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-10

Query Match 41.6%; Score 511.5; DB 4; Length 1151;
Best Local Similarity 33.0%; Pred. No. 7e-35;
Matches 94; Conservative 22; Mismatches 82; Indels 87; Gaps 3;

QY 3 HGGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSGLMRNC-QNKPCPDVG 61
DB 664 NGATPWSSWALCSTSGGIGFQVRQSCSNPAPRHGRICVKSREERCNENPCPVPI 723
QY 62 SWSLSWECTSCGRGNQTRTTCNNPSVQHGRCPCGNVAEIMCNIRPCP----- 115
DB 724 FWASGWSKSCSSCGGMSRRACEN-----GNSCLGCGVEFKTCNPEGCPVRNT 777
QY 116 ----- 115

DB 778 PWTPLPVNVTQGAQRQORFRTCTRAPLADPHLQFGRRRTTCTCPADGSGSDTDAL 837
QY 116 -----VHGASWOPWGTCSGKGTQTRARLNNPPPPAFGGSYCDGAET 161
DB 838 VEDLLRSGTSPHTVSGGAAWGPWSSCSRDCELGFRVRKKTCTNPEFRNGGLPCVGDAA 897
QY 162 QMVCNERNCPHGWKATWASWSACSVCSSGASQRTGCGSDPVP 206
DB 898 EYQDCNPQACVRGAWSCWTSWSPCSASCGGHHYQRTSCTSPAP 942

RESULT 9
US-08-985-526-3
; Sequence 3, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware

COUNTRY: U.S.A.
ZIP: 19899
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorrow Jr., Robert G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-985-526-3

Query Match 40.3%; Score 495.5; DB 3; Length 441;
Best Local Similarity 30.5%; Pred. No. 5.7e-34;
Matches 93; Conservative 31; Mismatches 72; Indels 109; Gaps 5;

QY 4 GGFQWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSGLMRNCQNKPCPDVGSW 63
DB 136 GGNHWSFWSSCVTCGDIVITITLCNPSQPMNGKPCGEARETRKACKDKACPIINGW 195
QY 64 SEMSLWECTSCGRGNQTRTTC-----NN----- 89
DB 196 GPNWSPWDICSVTCGGVQKRSRLCVDSRMTENKELANERPLCYHNGVQVRNNEWT 255
QY 90 -----PS----- 91

DB 256 VDSCTECHCONSVTICKKVCIPMPCSNATVPDGCPCRCWPSDSDADDGWSWSEWTSCS 315
QY 92 -----VQHGRP-----CEGNAVEIIMCNIRPCPVH-----GAKSAMOPWGTCSG 134
DB 316 TSCNGIQQRSCDSLNNRCSSSVQTRTCHIQEDCKRFKQGGHWSFWSSCVTCG 375
QY 135 KGTQTRARLNNPPPPAFGGSYCDGAETQMVQCNERNCPHGWKATWASWSACSVCSSGGA 194
DB 376 DGVITITLCNPSQPMNGKPCGEARETRKACKDKACPIINGWPNWSPWDICSVTCGGV 435
QY 195 RQRT 199
DB 436 QKRSR 440

RESULT 10
US-08-313-288B-20
; Sequence 20, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Query Match 38.9%; Score 478.5; DB 5; Length 239;
Best Local Similarity 46.9%;
Pred. No. 8.1e-33;

Patent No. 5876963
GENERAL INFORMATION:
APPLICANT: Mitchell, Peter
APPLICANT: Hutchinson, Nancy
APPLICANT: Lawton, Michael
APPLICANT: Magna, Holly
APPLICANT: Yocum, Sue
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,914
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0369
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 788 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1070094
US-08-918-914-4

Query Match 33.9%; Score 417.5; DB 2; Length 788;
Best Local Similarity 39.0%; Pred. No. 3.6e-27;
Matches 78; Conservative 20; Mismatches 79; Indels 23; Gaps 6;
QY 2 VHGGFQWSAWRACSVTCGKGIQKSRRLCNQPLPANGGKPCQGSDLNMRNCONKPCPVDG 61
Db 454 VSGVHDSWDSWTCSTCGDGAKSRRECEST-----NNCGADYETPCNLGFCQ--- 503
QY 62 SWSWSLWEECTRCGRGNQTRTCNNPNSVOHGRPCEGNAVEIMCNIRPCPVHGWS 121
Db 504 TWSWCEWSTCSASCGSGQRERTRFCH-----LGTNRCEGKDYSEQCSAGPCP---EWS 555
QY 122 ANQPWGTCSGSGKGTQTARLCNNPPAPFGGSDGAEOTMQVGNERNCPHKGWATWA 181
Db 556 QWEDWQCQSVTCQGVAVNQRCTCLG--GVFGHLCQGPKEACDGGPCSL---WSPNQ 610
QY 182 SWSACSVSCGGGARQTRGC 201
Db 611 EWSSTCSASCGSGMKRRQRC 630

Search completed: November 17, 2004, 15:57:03
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 15:52:32 ; Search time 140 seconds
(without alignments)
525.132 Million cell updates/sec

Title: US-10-019-065A-1

Perfect score: 1231

Sequence: 1 QVHGFSQWSAWRACSVTCG.....SCGGARQTRGCSDPVPQY 208

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.rep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCR_NEW_PUB.rep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.rep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.rep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.rep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.rep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.rep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.rep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.rep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.rep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.rep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.rep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.rep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.rep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.rep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.rep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.rep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.rep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.rep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.rep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231	100.0	208	15	US-10-019-065A-1
2	1231	100.0	707	15	US-10-138-588-18
3	1231	100.0	1336	15	US-10-019-065A-33
4	1231	100.0	1902	14	US-10-114-153-88
5	1231	100.0	3645	14	US-10-120-801-73
6	1231	100.0	5635	15	US-10-451-168-78
7	1231	100.0	5636	14	US-10-032-189-128
8	1231	100.0	5636	14	US-10-120-801-72
9	1231	100.0	5636	14	US-10-023-634-93
10	1231	100.0	5636	16	US-10-408-765A-1895
11	1230	99.9	2572	14	US-10-114-153-86
12	1230	99.9	2673	14	US-10-120-801-76
13	1230	99.9	3645	14	US-10-032-189-127

14	1230	99.9	4495	15	US-10-138-588-20	Sequence 20, Appl
15	1169	95.0	197	15	US-10-019-065A-31	Sequence 31, Appl
16	852	69.2	634	15	US-10-115-479-86	Sequence 86, Appl
17	648	52.6	1953	15	US-10-115-479-88	Sequence 88, Appl
18	644	52.3	877	9	US-09-764-898-200	Sequence 200, Appl
19	633	51.4	800	9	US-09-764-893-124	Sequence 124, Appl
20	633	51.4	800	9	US-09-764-853-800	Sequence 800, Appl
21	633	51.4	800	9	US-09-764-898-272	Sequence 272, Appl
22	633	51.4	800	14	US-10-073-865-124	Sequence 124, Appl
23	533.5	43.3	1077	14	US-10-190-115-43	Sequence 43, Appl
24	533.5	43.3	1077	15	US-10-369-072-43	Sequence 43, Appl
25	522.5	42.4	1074	14	US-10-190-115-44	Sequence 44, Appl
26	522.5	42.4	1074	15	US-10-369-072-44	Sequence 44, Appl
27	520.5	42.3	1074	14	US-10-190-115-41	Sequence 41, Appl
28	520.5	42.3	1074	15	US-10-369-072-41	Sequence 41, Appl
29	519.5	42.2	194	15	US-10-019-065A-32	Sequence 32, Appl
30	519.5	42.2	1584	14	US-10-225-567A-342	Sequence 342, Appl
31	511.5	41.6	401	14	US-10-190-115-116	Sequence 116, Appl
32	511.5	41.6	939	9	US-09-854-845-16	Sequence 16, Appl
33	511.5	41.6	954	9	US-09-854-845-14	Sequence 14, Appl
34	511.5	41.6	1034	9	US-09-854-845-6	Sequence 6, Appl
35	511.5	41.6	1049	9	US-09-854-845-2	Sequence 2, Appl
36	511.5	41.6	1078	9	US-09-854-845-8	Sequence 8, Appl
37	511.5	41.6	1093	9	US-09-854-845-4	Sequence 4, Appl
38	511.5	41.6	1093	14	US-10-102-524-1858	Sequence 1858, Ap
39	511.5	41.6	1093	14	US-10-190-115-106	Sequence 106, Appl
40	511.5	41.6	1130	14	US-10-190-115-108	Sequence 108, Appl
41	511.5	41.6	1136	9	US-09-854-845-12	Sequence 12, Appl
42	511.5	41.6	1151	9	US-09-854-845-10	Sequence 10, Appl
43	511.5	41.6	1151	14	US-10-102-524-1857	Sequence 1857, Ap
44	511.5	41.6	1202	14	US-10-102-524-1856	Sequence 1856, Ap
45	511.5	41.6	1202	14	US-10-190-115-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-10-019-065A-1
; Sequence 1, Application US/10019065A
; Publication No. US20040086501A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Corporation
; TITLE OF INVENTION: Protein Having Activity As An Angiogenesis Modulator
; FILE REFERENCE: MSB-7265-PCT
; CURRENT APPLICATION NUMBER: US/10/019,065A
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/266,300
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-019-065A-1

Query Match	100.0%	Score 1231;	DB 15;	Length 208;
Best Local Similarity	100.0%	Pred. No. 7.6e-93;		
Matches 208;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	QVHGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGS	DLERNCNKPCFVD	60
Db	1	QVHGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGS	DLERNCNKPCFVD	60
Qy	61	GSWSESLWBECTSCRGNGQTRTRTCNNPSVQHGGRPCGNAVEIIMCNRIRPCFVHGA	W 120	
Db	61	GSWSESLWBECTSCRGNGQTRTRTCNNPSVQHGGRPCGNAVEIIMCNRIRPCFVHGA	W 120	
Qy	121	SAWQPMGTCSBSCGKGIQTRARLNCNNPPAFGSGYCDGAETQMCVNCNRCPIHGKWA	T 180	
Db	121	SAWQPMGTCSBSCGKGIQTRARLNCNNPPAFGSGYCDGAETQMCVNCNRCPIHGKWA	T 180	

Qy	181	ASWSACSVSCGGGARQTRGCDPVPQY	208
Db	181	ASWSACSVSCGGGARQTRGCDPVPQY	208


```
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/283710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/283678
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284234
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 251
; SEQ ID NO 88
; LENGTH: 1902
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-153-88

Query Match      100.0%; Score 1231; DB 14; Length 1902;
Best Local Similarity 100.0%; Pred. No. 5.9e-92;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSDEMRNCQNKPCPD 60
Db 911 QVHGGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSDEMRNCQNKPCPD 970

QY 61 GSWSESLWBEETRSCGRGNQTRTRTCNNPSPVQHGGRPCGNAVEIIMCNIIRPCPVHGAW 120
Db 971 GSWSESLWBEETRSCGRGNQTRTRTCNNPSPVQHGGRPCGNAVEIIMCNIIRPCPVHGAW 1030

QY 121 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETOMQVCNERNCPHKGKATW 180
Db 1031 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETOMQVCNERNCPHKGKATW 1090

QY 181 ASWSACSVSCGGGARQTRGCDPVPQY 208
Db 1091 ASWSACSVSCGGGARQTRGCDPVPQY 1118

RESULT 5
US-10-120-801-73
; Sequence 73, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCES: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
```

```
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 3645
; TYPE: PRT
; ORGANISM: human
US-10-120-801-73

Query Match      100.0%; Score 1231; DB 14; Length 3645;
Best Local Similarity 100.0%; Pred. No. 1.1e-91;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSDEMRNCQNKPCPD 60
Db 2537 QVHGGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSDEMRNCQNKPCPD 2596

QY 61 GSWSESLWBEETRSCGRGNQTRTRTCNNPSPVQHGGRPCGNAVEIIMCNIIRPCPVHGAW 120
Db 2597 GSWSESLWBEETRSCGRGNQTRTRTCNNPSPVQHGGRPCGNAVEIIMCNIIRPCPVHGAW 2656

QY 121 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETOMQVCNERNCPHKGKATW 180
Db 2657 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETOMQVCNERNCPHKGKATW 2716

QY 181 ASWSACSVSCGGGARQTRGCDPVPQY 208
Db 2717 ASWSACSVSCGGGARQTRGCDPVPQY 2744

RESULT 6
US-10-451-168-78
; Sequence 78, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 5635
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-78

Query Match      100.0%; Score 1231; DB 15; Length 5635;
Best Local Similarity 100.0%; Pred. No. 1.6e-91;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGGKPCQGSDEMRNCQNKPCPD 60
```

Db 4527 QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSDEMRNCONKPCPVD 4586
 QY 61 GSWSEWSLWECTRSCGRNQTRTTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGAW 120
 Db 4587 GSWSEWSLWECTRSCGRNQTRTTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGAW 4646
 QY 121 SAWPWTCTSCSCGKGTQTRARLNCNNPPAFGSGYCDGAEQTMQVNCNERNCPHGWATW 180
 Db 4647 SAWPWTCTSCSCGKGTQTRARLNCNNPPAFGSGYCDGAEQTMQVNCNERNCPHGWATW 4706
 QY 181 ASWSACSVSCGGARQTRGSDVPVQY 208
 Db 4707 ASWSACSVSCGGARQTRGSDVPVQY 4734

RESULT 7

US-10-032-189-128
 ; Sequence 128, Application US/10032189
 ; Publication No. US20030170630A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alsbrook II, John P
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Zerhusen, Bryan D
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Lepley, Denise M
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Szekeres, Edward S
 ; APPLICANT: Vernet, Corine A.M.
 ; APPLICANT: Li, Li
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Boldog, Ferenc L
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Fernandes, Elma R
 ; APPLICANT: Rieger, Daniel K
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: MacDougall, John R
 ; APPLICANT: Smithson, Glenda
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-228
 ; CURRENT APPLICATION NUMBER: US/10/032,189
 ; CURRENT FILING DATE: 2001-12-21
 ; PRIOR APPLICATION NUMBER: 60/257,495
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/258,171
 ; PRIOR FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 60/269,940
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/274,192
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/277,826
 ; PRIOR FILING DATE: 2001-03-22
 ; PRIOR APPLICATION NUMBER: 60/279,840
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/282,981
 ; PRIOR FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 60/283,656
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: 60/309,247
 ; PRIOR FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: 60/311,754
 ; PRIOR FILING DATE: 2001-08-17
 ; PRIOR APPLICATION NUMBER: 60/313,331
 ; PRIOR FILING DATE: 2001-08-17

; NUMBER OF SEQ ID NOS: 260
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 128
 ; LENGTH: 5636
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (3003)
 ; OTHER INFORMATION: Wherein Xaa is any amino acid.
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (3041)
 ; OTHER INFORMATION: Wherein Xaa is any amino acid.
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (3367)
 ; OTHER INFORMATION: Wherein Xaa is any amino acid.
 ; US-10-032-189-128
 Query Match 100.0%; Score 1231; DB 14; Length 5636;
 Best Local Similarity 100.0%; Pred. No. 1.6e-91;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSDEMRNCONKPCPVD 60
 Db 4528 QVHGFSQSAWRACSVTCGKGIQKRSRLCNQPLPANGKPCQGSDEMRNCONKPCPVD 4587
 QY 61 GSWSEWSLWECTRSCGRNQTRTTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGAW 120
 Db 4588 GSWSEWSLWECTRSCGRNQTRTTCNNPVSQHGGRPCGEGNAVEIIMCNIRPCPVHGAW 4647
 QY 121 SAWPWTCTSCSCGKGTQTRARLNCNNPPAFGSGYCDGAEQTMQVNCNERNCPHGWATW 180
 Db 4648 SAWPWTCTSCSCGKGTQTRARLNCNNPPAFGSGYCDGAEQTMQVNCNERNCPHGWATW 4707
 QY 181 ASWSACSVSCGGARQTRGSDVPVQY 208
 Db 4708 ASWSACSVSCGGARQTRGSDVPVQY 4735
 RESULT 8
 US-10-120-801-72
 ; Sequence 72, Application US/10120801
 ; Publication No. US20030203843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pena, Carol
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Shimkets, Richard
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Mehraban, Fuad
 ; APPLICANT: Topper, James N.
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Wasserman, Scott
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Komuves, Laszlo
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-340
 ; CURRENT APPLICATION NUMBER: US/10/120,801
 ; CURRENT FILING DATE: 2002-04-11
 ; PRIOR APPLICATION NUMBER: 60/285748
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: 60/286068
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/286292
 ; PRIOR FILING DATE: 2001-04-25
 ; PRIOR APPLICATION NUMBER: 60/288334
 ; PRIOR FILING DATE: 2001-05-03
 ; PRIOR APPLICATION NUMBER: 60/291241

```
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-221
; CURRENT APPLICATION NUMBER: US/10/023,634
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,025
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: 60/265,163
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,929
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/274,864
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/276,688
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,880
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/286,409
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/309,246
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/315,600
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 5636
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (3003)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification.
; NAME/KEY: VARIANT
; LOCATION: (3041)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification.
; NAME/KEY: VARIANT
; LOCATION: (3367)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification.
; US-10-023-634-93

Query Match      100.0%; Score 1231; DB 14; Length 5636;
Best Local Similarity 100.0%; Pred. No. 1.6e-91;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKCGSDLEMRNCQKPCPVD 60
DB 4528 QVHGFSQWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKCGSDLEMRNCQKPCPVD 4587
QY 61 GSSEWSLWEECTRS CGRGNQTRTRTCNNPSVQHGRPCGEGNAVEIIMCNIIRPCPVHGW 120
DB 4588 GSSEWSLWEECTRS CGRGNQTRTRTCNNPSVQHGRPCGEGNAVEIIMCNIIRPCPVHGW 4647
QY 121 SAWQPMGTCSGCGKTQTRARLCNNPPAPFGSYCDGAETQMVQCNERNCPHKGKATW 180
DB 4648 SAWQPMGTCSGCGKTQTRARLCNNPPAPFGSYCDGAETQMVQCNERNCPHKGKATW 4707
QY 181 ASWSACSVSCGGGARQRTGCDPVPQY 208
DB 4708 ASWSACSVSCGGGARQRTGCDPVPQY 4735

RESULT 10
US-10-408-765A-1895
; Sequence 1895, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
```

```

; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1895
; LENGTH: 5636
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 3003, 3041, 3367
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-408-765A-1895

Query Match 100.0%; Score 1231; DB 16; Length 5636;
Best Local Similarity 100.0%; Pred. No. 1.6e-91;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVHGGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMRNCKPCPD 60
Db 4528 QVHGGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMRNCKPCPD 4587

Qy 61 GSWSESLWEECTSCRGNGQTTRICNNPSVOHGGPCEGNAVEIMCNIIRPCPVHGA 120
Db 4588 GSWSESLWEECTSCRGNGQTTRICNNPSVOHGGPCEGNAVEIMCNIIRPCPVHGA 4647

Qy 121 SAWQPGTCSGCKGTQTRARLNNPPPAFGGSYCDGAEQVQCNERNCPHGWATW 180
Db 4648 SAWQPGTCSGCKGTQTRARLNNPPPAFGGSYCDGAEQVQCNERNCPHGWATW 4707

Qy 181 ASWSACSVSCGGARQTRGCSDFVPQY 208
Db 4708 ASWSACSVSCGGARQTRGCSDFVPQY 4735

RESULT 11
US-10-114-153-86
; Sequence 86, Application US/10114153
; Publication No. US20030185815A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glenda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Spytek, Kimberly
; APPLICANT: Malyankar, Uriel
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Heyes, Melvyn
; APPLICANT: Ju, Jingfang
; APPLICANT: Peyman, John
; APPLICANT: Catterton, Elina

; APPLICANT: MacDougall, John
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Mazur, Ann
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACID
; FILE REFERENCE: 21402-322A
; CURRENT APPLICATION NUMBER: US/10/114,153
; CURRENT FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: 60/281086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283444
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283657
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/283710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/283678
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284234
; PRIOR FILING DATE: 2001-04-17
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 251
; SEQ ID NO 86
; LENGTH: 2572
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-153-86

Query Match 99.9%; Score 1230; DB 14; Length 2572;
Best Local Similarity 99.5%; Pred. No. 9.4e-92;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVHGGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMRNCKPCPD 60
Db 1581 QVHGGFSQSAWRACSVTCGKGIQKSRCLCNQPLPANGKPKCGSDLEMRNCKPCPD 1640

Qy 61 GSWSESLWEECTSCRGNGQTTRICNNPSVOHGGPCEGNAVEIMCNIIRPCPVHGA 120
Db 1641 GSWSESLWEECTSCRGNGQTTRICNNPSVOHGGPCEGNAVEIMCNIIRPCPVHGA 1700

Qy 121 SAWQPGTCSGCKGTQTRARLNNPPPAFGGSYCDGAEQVQCNERNCPHGWATW 180
Db 1701 SAWQPGTCSGCKGTQTRARLNNPPPAFGGSYCDGAEQVQCNERNCPHGWATW 1760

Qy 181 ASWSACSVSCGGARQTRGCSDFVPQY 208
Db 1761 ASWSACSVSCGGARQTRGCSDFVPQY 1788

RESULT 12
US-10-120-801-76
; Sequence 76, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehra, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott

```

APPLICANT: Edinger, Shlomit
APPLICANT: Smithson, Glennda
APPLICANT: Gunther, Erik
APPLICANT: Komuves, Laszlo
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-340
CURRENT APPLICATION NUMBER: US/10/120,801
CURRENT FILING DATE: 2002-04-11
PRIOR APPLICATION NUMBER: 60/285748
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 60/286068
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/286292
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/288334
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/291241
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/322284
PRIOR FILING DATE: 2001-09-14
PRIOR APPLICATION NUMBER: 60/285609
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 155
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 76
LENGTH: 2673
TYPE: PRT
ORGANISM: human
FEATURE:
NAME/KEY: VARIANT
LOCATION: (119)
OTHER INFORMATION: Wherein Xaa is any amino acid.

US-10-120-801-76

Query Match 99.9%; Score 1230; DB 14; Length 2673;
Best Local Similarity 99.5%; Pred. No. 9.7e-92; Indels 0; Gaps 0;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKQGSDELNRNCQKPCPVD 60
Db 1565 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKQGSDELNRNCQKPCPVD 1624
QY 61 GSSEWSLWEETRSCGRGNQTRTCNNPVSVOHGRPCGNAVEIIMCNIRPCPVHGA 120
Db 1625 GSSEWSLWEETRSCGRGNQTRTCNNPVSVOHGRPCGNAVEIIMCNIRPCPVHGA 1684
QY 121 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHKGKATW 180
Db 1685 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHKGKATW 1744
QY 181 ASWSACSVSCGGGARQTRGCDPVPQY 208
Db 1745 ASWSACSVSCGGGARQTRGCDPVPQY 1772

RESULT 13
US-10-032-189-127
Sequence 127, Application US/10032189
Publication No. US20030170630A1
GENERAL INFORMATION:
APPLICANT: Alsbrook II, John P
APPLICANT: Tchernev, Velizar T
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly A
APPLICANT: Zerhusen, Bryan D
APPLICANT: Patturajan, Meera
APPLICANT: Grosse, William M
APPLICANT: Lepley, Denise M
APPLICANT: Burgess, Catherine B
APPLICANT: Shimkets, Richard A
APPLICANT: Grosse, William M
APPLICANT: Szekeres, Edward S
APPLICANT: Vernet, Corine A.M.

APPLICANT: Li, Li
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc L
APPLICANT: Gorman, Linda
APPLICANT: Gangolli, Esha A
APPLICANT: Fernandes, Elma R
APPLICANT: Rieger, Daniel K
APPLICANT: Edinger, Shlomit R
APPLICANT: Gunther, Erik
APPLICANT: Millet, Isabelle
APPLICANT: Sciore, Paul
APPLICANT: Ellerman, Karen
APPLICANT: MacDougall, John R
APPLICANT: Smithson, Glennda
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-228
CURRENT APPLICATION NUMBER: US/10/032,189
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 60/257,495
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/258,171
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 60/269,940
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/274,192
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/277,826
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 60/279,840
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: 60/282,981
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 60/283,656
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/309,247
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/311,754
PRIOR FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: 60/313,331
PRIOR FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 260
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 127
LENGTH: 3645
TYPE: PRT
ORGANISM: Homo sapiens
US-10-032-189-127
Query Match 99.9%; Score 1230; DB 14; Length 3645;
Best Local Similarity 99.5%; Pred. No. 1.3e-91;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKQGSDELNRNCQKPCPVD 60
Db 2537 QVHGFQSWSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKQGSDELNRNCQKPCPVD 2596
QY 61 GSSEWSLWEETRSCGRGNQTRTCNNPVSVOHGRPCGNAVEIIMCNIRPCPVHGA 120
Db 2597 GSSEWSLWEETRSCGRGNQTRTCNNPVSVOHGRPCGNAVEIIMCNIRPCPVHGA 2656
QY 121 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHKGKATW 180
Db 2657 SAWQPMGTCSGCKGTQTRARLCNNPPAFGGSYCDGAETQMVQCNERNCPHKGKATW 2716
QY 181 ASWSACSVSCGGGARQTRGCDPVPQY 208
Db 2717 ASWSACSVSCGGGARQTRGCDPVPQY 2744
RESULT 14
US-10-138-588-20
Sequence 20, Application US/10138588
Publication No. US20040018594A1

```

; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: NOVEL ANTIBODIES THAT BIND TO ANTIGENIC POLYPEPTIDES, NUCLEIC ACID SEQUENCES, AND METHODS OF USE
; FILE REFERENCE: 21402-347A
; CURRENT APPLICATION NUMBER: US/10/138,588
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/288,395
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/308,901
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/313,388
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/324,757
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/288,900
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 203
; SEQ ID NO 20
; LENGTH: 4495
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-138-588-20

Query Match          99.9%; Score 1230; DB 15; Length 4495;
Best Local Similarity 99.5%; Pred. No. 1.6e-91;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQSAWRACSVTCCKGKIQKSRSLCNQPLPANGKPKCGSDLEVRNCKPCPYD 60
DB 3387 QVHGFSQSAWRACSVTCCKGKIQKSRSLCNQPLPANGKPKCGSDLEVRNCKPCPYD 3446

QY 61 GSWSESLWEECTRCGRGNQTRTRCNPNPSVQHGGRPCEGNAVEIIMCNIRPCPVHGW 120
DB 3447 GSWSESLWEECTRCGRGNQTRTRCNPNPSVQHGGRPCEGNAVEIIMCNIRPCPVHGW 3506

QY 121 SAWQPMGTSCSCGKGTQTRARLCNNPPPAFGGSYCDGAETQMCVNCERNCPHKGKATW 180
DB 3507 SAWQPMGTSCSCGKGTQTRARLCNNPPPAFGGSYCDGAETQMCVNCERNCPHKGKATW 3566

QY 181 ASWSACSVSCGGARQTRCGSDPPQY 208
DB 3567 ASWSACSVSCGGARQTRCGSDPPQY 3594

RESULT 15
US-10-019-065A-31
; Sequence 31, Application US/10019065A
; Publication No. US20040086501A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Corporation
; TITLE OF INVENTION: Protein Having Activity As An Angiogenesis Modulator
; FILE REFERENCE: MSB-7265-PCT
; CURRENT APPLICATION NUMBER: US/10/019,065A
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 60/266,300
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-019-065A-31

Query Match          95.0%; Score 1169; DB 15; Length 197;
Best Local Similarity 100.0%; Pred. No. 8.3e-88;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 QWSAWRACSVTCCKGKIQKSRSLCNQPLPANGKPKCGSDLEVRNCKPCPYDGSWSWS 67
DB 1 QWSAWRACSVTCCKGKIQKSRSLCNQPLPANGKPKCGSDLEVRNCKPCPYDGSWSWS 60

```

```

QY 68 LWECTRCGRGNQTRTRCNPNPSVQHGGRPCEGNAVEIIMCNIRPCPVHGWSAWQFWG 127
DB 61 LWECTRCGRGNQTRTRCNPNPSVQHGGRPCEGNAVEIIMCNIRPCPVHGWSAWQFWG 120

QY 128 TCSESCGKGTQTRARLCNNPPPAFGGSYCDGAETQMCVNCERNCPHKGKATWASWSACS 187
DB 121 TCSESCGKGTQTRARLCNNPPPAFGGSYCDGAETQMCVNCERNCPHKGKATWASWSACS 180

QY 188 VSCGGARQTRCGSDP 204
DB 181 VSCGGARQTRCGSDP 197

```

Search completed: November 17, 2004, 16:07:23
Job time : 141 secs

brain-specific angiogenesis inhibitor 1 - human

N:Alternate names: BAI1 protein

C:Species: Homo sapiens (man)

C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C:Accession: T00026

R:Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida

submitted to the EMBL Data Library, June 1997

A:Reference number: 214064

A:Accession: T00026

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1584 <NIS>

A:Cross-references: UNIPROT:O14514; EMBL:AB005297; NID:d1175078; PID:d1024528

A:Experimental source: brain

C:Genetics:

A:Gene: GDB:BAI1

A:Cross-references: GDB:9838088; OMIM:602682

A:Map position: 8q24-8q24

F:408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 42.2%; Score 519.5; DB 2; Length 1584;

Best Local Similarity 45.8%; Pred. No. 2.3e-31;

Matches 92; Conservative 25; Mismatches 73; Indels 11; Gaps 5;

QY 8 QWSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSN 66

DB 359 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 415

QY 67 SLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCP---VHGAWSAW 123

DB 416 SPWLSLSTCGRFRDRTTRCP--QFGNGPCGPKQTFKFNIALCPRAVDGNW 473

QY 124 QPWGTCSESCGKGTQTRARLCNNPPAFGSGYCDGAETQMCVNERNCPIHGKWATWASW 183

DB 474 SSWACASCSQGRQRTRECN--PSYGAECQGHVETRDCLFQCCPVDGKQWQW 531

QY 184 SACSVCSCGGARQTRGSDP 204

DB 532 GSCSVTCGAGSQRREVCSPG 552

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 415

QY 124 QPWGTCSESCGKGTQTRARLCNNPPAFGSGYCDGAETQMCVNERNCPIHGKWATWASW 183

DB 474 SSWACASCSQGRQRTRECN--PSYGAECQGHVETRDCLFQCCPVDGKQWQW 531

QY 184 SACSVCSCGGARQTRGSDP 205

DB 465 SGCSKSCDGGWERRIRTCQGA 486

Query Match 40.9%; Score 503.5; DB 2; Length 984;

Best Local Similarity 46.0%; Pred. No. 2.4e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 406

QY 124 QPWGTCSESCGKGTQTRARLCNNPPAFGSGYCDGAETQMCVNERNCPIHGKWATWASW 183

DB 407 SWSQCSVTCNGTQQRROCT--AAAHGSGECRGPWAESRECYNPECTANGQWQWGH 464

QY 184 SACSVCSCGGARQTRGSDP 205

DB 465 SGCSKSCDGGWERRIRTCQGA 486

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 415

QY 124 QPWGTCSESCGKGTQTRARLCNNPPAFGSGYCDGAETQMCVNERNCPIHGKWATWASW 183

DB 407 SWSQCSVTCNGTQQRROCT--AAAHGSGECRGPWAESRECYNPECTANGQWQWGH 464

QY 184 SACSVCSCGGARQTRGSDP 205

DB 465 SGCSKSCDGGWERRIRTCQGA 486

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 415

QY 124 QPWGTCSESCGKGTQTRARLCNNPPAFGSGYCDGAETQMCVNERNCPIHGKWATWASW 183

DB 407 SWSQCSVTCNGTQQRROCT--AAAHGSGECRGPWAESRECYNPECTANGQWQWGH 464

QY 184 SACSVCSCGGARQTRGSDP 205

DB 465 SGCSKSCDGGWERRIRTCQGA 486

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 415

QY 124 QPWGTCSESCGKGTQTRARLCNNPPAFGSGYCDGAETQMCVNERNCPIHGKWATWASW 183

DB 407 SWSQCSVTCNGTQQRROCT--AAAHGSGECRGPWAESRECYNPECTANGQWQWGH 464

QY 184 SACSVCSCGGARQTRGSDP 205

DB 465 SGCSKSCDGGWERRIRTCQGA 486

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 406

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

DB 293 GVEEWSQWSTCSVTCCGSGQVTRTCVSPY---GTHCSGPLRESRVNNTALCPVHG 348

QY 64 SEWSLWEECTSCRGNGQTRTRCNPNPSVQHGGRPCGNAVEIIMCNIRPCPVGAWSAW 123

DB 349 EWSFWSVCSSTCGEGWQTRTRFC---VSSSVSTQCSGPLREQLCNNSAVCPVHGAWDEW 406

Query Match 40.9%; Score 503.5; DB 2; Length 1522;

Best Local Similarity 46.0%; Pred. No. 3.5e-30;

Matches 93; Conservative 15; Mismatches 85; Indels 9; Gaps 4;

QY 5 GFSQMSAWRACSVTCGKGIQKRSLCNQPLPANGKPCQGSQSDLEMRNCQKP-CPVDGWSW 63

A:Accession: A37905
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <BOR>
A:Cross-references: GB:J05605; GB:J05606; NID:G201991; PIDN:AAA40431.1; PID:G554390
R:Lawler, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seidlin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: B42587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1152, 'P', 1154-1170 <LAH>
A:Cross-references: GB:M87276
A:Note: sequence extracted from NCBI backbone (NCBIP:81501)
R:Chen, H.; Aeschlimann, D.; Nowlin, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A:Title: Expression and initial characterization of recombinant mouse thrombospondin 1 a
A:Reference number: S68787; MUID:96234006; PMID:8654563
A:Accession: S68787
A:Molecule type: protein
A:Residues: 19-26, 'X', 28-37 <CHE>
C:Complex: homotrimer, disulfide linked
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C:Keywords: calcium binding; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1170/Product: thrombospondin 1 #status predicted <MAT>
F:317-375/Domain: von Willebrand factor type C repeat homology <VMC>
F:378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F:551-586/Domain: EGF homology <EGF>
F:248,360,708,1067/Binding site: carbohydrate (Asn) #status predicted
Query Match 40.0%; Score 492.5; DB 2; Length 1170;
Best Local Similarity 42.9%; Pred. No. 1.9e-29;
Matches 93; Conservative 23; Mismatches 74; Indels 27; Gaps 6;
QY 5 GFQSWARACSVTCGKGIQKRSRLCNQPLPANGKPGCGSDLEMRNCQKPC---PVD 60
DB 381 GWSPESEWTSCTATCGNGIQGRSCD----SLNRCBSSVQRTTCHIQCDKRFKQD 435
QY 61 GSWSEVSLWBEETRSCGRGNQTRTCNNPSPVHGRRPCEGNAVEIIMCNIRPCPVHGAW 120
DB 436 GWSHSPSSSCSVTCGDDGVTIRILCNPSFPQMGKPCGEARETKACKDACPINGW 495
QY 121 SAMQPGTCSSECGKGTQTRARLCNNPPPAFGSGYCDGAETQMVQNCERNCPHIG---- 175
DB 496 GPNSPWDICSVTCGGVQRSLRNLNPTTFQGGKQCVGVNTENQVCKQDCPDIDGLSNP 555
QY 176 -----KWATW--ASW--SACSVSCGGGARQRTRGCS 203
DB 556 CFAGAKTSPYDGSWKGACGPGYSGNGIQ-----CKD 588
RESULT 6
T00027
N:Alternative names: BAI2 protein
C:Species: Homo sapiens (man)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C:Accession: T00027
R:Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.
Cytogenet. Cell Genet. 79, 103-108, 1997
A:Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-
A:Reference number: Z14066; MUID:98194217; PMID:9533023
A:Accession: T00027
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1572 <SHI>
A:Cross-references: UNIPROT:O60241; EMBL:AB005298; NID:G3021698; PIDN:BA2A25362.1; PID:G3
A:Experimental source: brain
C:Genetics:

A:Gene: GDB:BAI2
A:Cross-references: GDB:9838089; OMIM:602683
A:Map position: lp35-lp35
Query Match 39.8%; Score 489.5; DB 2; Length 1572;
Best Local Similarity 44.9%; Pred. No. 3.9e-29;
Matches 88; Conservative 27; Mismatches 72; Indels 9; Gaps 5;
QY 8 QWSARACSVTCGKGIQKRSRLCNQPLPANGKPGCGSDLEMRNCQK-PCPVHGSAWSEW 66
DB 302 EWSPEVSLWBEETRSCGRGNQTRTCNNPSPVHGRRPCEGNAVEIIMCNIRPCPVHGSAWSEW 358
QY 67 SLWBEETRSCGRGNQTRTCNNPSPVHGRRPCEGNAVEIIMCNIRPCPVHGSAWSEW 126
DB 359 GSWSLRSCGRGSRSRMRTCVPP--QGGKACBGPPLQTKLSMAACPFVGGWLEHGPW 416
QY 127 QTCSECGKGTQTRARLCNNPPPAFGSGYCDGAETQMVQNCERNCP-IGHKWATWASWSA 185
DB 417 GPCSTSCANGTQQRSRKCSVAGPAW--ATCTGALTDTRECSNLECPATDSKGGPWNWSL 474
QY 186 CSVSCGGGARQRTGCG 201
DB 475 CSKTCDDTGWQRRFRWC 490
RESULT 7
TSHUP1
Thrombospondin 1 precursor - human
C:Species: Homo sapiens (man)
C:Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: A26155; A34274; A30140; A25812; A05172; A42927
R:Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple c
A:Reference number: A26155; MUID:87057617; PMID:2430973
A:Accession: A26155
A:Molecule type: mRNA
A:Residues: 1-1170 <LAH>
A:Cross-references: UNIPROT:P07996; GB:X04665; NID:G37137; PIDN:CAA28370.1; PID:G37138
A:Note: parts of this sequence, including the amino end of the mature protein, were dete
R:Lawler, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA s
A:Reference number: A34274; MUID:89291870; PMID:2544587
A:Accession: A34274
A:Molecule type: DNA
A:Residues: 1-166 <LAH>
A:Cross-references: GB:J04835
R:Hennessey, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, J.
J. Cell Biol. 108, 729-736, 1989
A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the
A:Reference number: A30140; MUID:89139590; PMID:2918029
A:Accession: A30140
A:Molecule type: mRNA
A:Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>
A:Cross-references: EMBL:X14787; NID:G37464; PIDN:CAA32889.1; PID:G37465
A:Note: parts of this sequence, including the amino end of the mature protein, were dete
R:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A:Title: Partial amino acid sequence of human thrombospondin as determined by analysis o
A:Reference number: A25812; MUID:87157592; PMID:3030396
A:Accession: A25812
A:Molecule type: mRNA
A:Residues: 1-83, 'A', 85-397 <KOB>
A:Cross-references: GB:M25631; NID:G538353; PIDN:AAA36741.1; PID:G538354
R:Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A:Reference number: A05172; MUID:86287276; PMID:3461443
A:Accession: A05172
A:Molecule type: mRNA
A:Residues: 1-83, 'A', 85-374, 'RC' <DIX>
A:Cross-references: GB:M14326; NID:G340005; PIDN:AAA61237.1; PID:G553801
A:Note: parts of this sequence, including the amino end of the mature protein, were dete

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A:Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A:Reference number: A42927; MUID:92348511; PMID:1379247
A:Accession: A42927
A:Molecule type: protein
A:Residues: 987-1003 <SUN>
A:Note: Cys-992 is shown to have a free sulphydryl
C:Genetics:
A:Gene: GDB:THS1; TSP1; TSP
A:Cross-references: GDB:120438; OMIM:188060
A:Map position: 15q15-15q15
A:Introns: 23/1
A:Note: the list of introns may be incomplete
C:Complex: homotrimer, disulfide linked
C:Function:
A:Description: participates in cell migration and adhesion, and in platelet aggregation
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1170/Product: thrombospondin 1 #status predicted <MAT>
F:317-375/Domain: von Willebrand factor type C repeat homology <WVC>
F:378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F:551-586/Domain: EGF homology <EGF1>
F:650-689/Domain: EGF homology <EGF2>
F:926-928/Region: cell attachment (R-G-D) motif
F:171-232/disulfide bonds: #status predicted
F:248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:270,274/disulfide bonds: interchain #status predicted
F:610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 39.7%; Score 488.5; DB 1; Length 1170;
Best Local Similarity 41.9%; Pred. No. 3.7e-29;
Matches 91; Conservative 26; Mismatches 73; Indels 27; Gaps 6;

QY 5 GFSSAWRACSVTCGKGIQKRLCNQPLPANGKPCQGSLEMRNCQKPC---PVD 60
Db 381 GWSFMSWTSCTSCGNGIQQRGSCD-----SLNRCGSSVQTRCHIQECDKRFKQD 435

QY 61 GSWSESLWEETRSCGRNQTRTRCNPNPSVQHGGPCGEGNAVEIMCNIRPCPVHGAW 120
Db 436 GWSHNSFWSCSVTCGVDGVTIRLNCSPSPQWNGKPCGEARETKACKKDACPINGGW 495

QY 121 SAWOPWGTCTSCSCGKGTQTRALCNPNPPAFGGSYCDGAEQVQVNERNCPIHG----- 175
Db 496 GWSFMDICSVTCGGVQKSRSLCNNTPOFGKDCVGDVTENQICNKQDCPIDGCLSNP 555

QY 176 -----KWATW--ASW--SACSVSCGGGARQTRGCSGD 203
Db 556 CFAGVKCTSYPDGSKGACGPPGYSNGIQ-----CTD 588

RESULT 8
S05478
properdin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S05478
R:Goundis, D.; Reid, K.B.M.
Nature 335, 82-85, 1988
A:Title: Properdin, the terminal complement components, thrombospondin and the circumpo
A:Reference number: S05478; MUID:88318954; PMID:3045564
A:Accession: S05478
A:Molecule type: mRNA
A:Residues: 1-437 <GOU>
A:Cross-references: UNIPROT:P11680; EMBL:X12905; NID:953786; PIDN:CAA31389.1; PID:953787
C:Complex: a mixture of homodimers, homotrimers and homotetramers
C:Function:
A:Description: protects C3 convertase (C3bBb) from rapid inactivation
A:Pathway: complement alternate pathway

C:Superfamily: human properdin precursor; thrombospondin type 1 repeat homology
C:Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer; homotrin
F:45-97/Domain: thrombospondin type 1 repeat homology <THR1>
F:104-160/Domain: thrombospondin type 1 repeat homology <THR2>
F:161-224/Domain: thrombospondin type 1 repeat homology <THR3>
F:225-282/Domain: thrombospondin type 1 repeat homology <THR4>
F:283-345/Domain: thrombospondin type 1 repeat homology <THR5>
F:346-408/Domain: thrombospondin type 1 repeat homology <THR6>
F:52,55,108,111,114,165,168,229,232,290,293,350,353,356/Modified site: 2'-mannosyl-tryptid
F:366,396/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 38.5%; Score 474.5; DB 2; Length 437;
Best Local Similarity 40.7%; Pred. No. 1.9e-28;
Matches 88; Conservative 30; Mismatches 81; Indels 17; Gaps 8;

QY 4 GGFSSAWRACSVTCGKGIQKRLCNQPLPANGKPCQGSLEMRNCQ-NKPCPVDS 62
Db 106 GWSSEWGPWPCSVTCGKGIQKRLCNQPLPANGKPCQGSLEMRNCQ-NKPCPVDS 164

QY 63 WSESLWEETRSCGRNQ-----TRTRCNPNPSVQH--GGRPCGEGNAVEIMCN-IRPCP 115
Db 165 WASNGPWSFSSGSLGGAQEPKTRSRSCSAPAFSHQPPGKPCGPAVEHKACSGLPCCP 224

QY 116 VHGSAWQPNWGTCTSCSCGKGTQTRALCNPNPPAFGGSYCDGAEQVQVNER-NCPIH 174
Db 225 VAGGWPWSPLSPCSVTGCGTLEQRTCDHPAPRHGGPFCAGDATRQMCNKAVPCFVN 284

QY 175 GKWATWASWSAC-----SVSCGG--GARQTRGCSGD 203
Db 285 GEWEANGWSDCSLRLWSINCEGTPGQSRSRSCGD 320

RESULT 9
T18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18856; T24653
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19031
A:Accession: T18856
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WIL>
A:Cross-references: UNIPROT:Q8MYA8; EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:CO:
A:Experimental source: clone C02B4
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19917
A:Accession: T24653
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <W12>
A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:CO2B4.1
A:Experimental source: clone T07C5
C:Genetics:
A:Gene: CESP:CO2B4.1
A:Map position: X
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566/

Query Match 38.3%; Score 472; DB 2; Length 1444;
Best Local Similarity 41.7%; Pred. No. 7.4e-28;
Matches 88; Conservative 25; Mismatches 86; Indels 12; Gaps 7;

QY 4 GGFSSAWRACSVTCG-KGQKSRSLCNQPLPANGKPCQGSLEMRNCQ-NKPC---P 58
Db 1187 GWSLWSEWSSCKDCGDTGTHQIRNRCSEPIPSNRGAYCSGYFQDQPCVMDNVCSDEK 1246

QY 59 VDGWSESLWEETRSCGRNQTRTRCNPNPSVQHGGPCGEGNAVEIMCNIRPCPVH- 117
Db 1247 VDGWTDWTAWSECTDYCRNGHRSRTFCANPKPSQGAQCTGSDFLNFC-FPPARCHL 1305

```
QY 118 --GAWSAWPGWTCSECGKGTQTRARLCNNPPPAFGSGYCDGAETOMQVNCERNIC--PI 173
DB 1306 RRGWSTWSDWTPCASCGFGVQTRDRSCSPEPK-GGQCSGLAHQISLCLDLPACDHES 1364

QY 174 HKWATWASWACSVCSSGGAARQTRGCSDP 204
DB 1365 DGWSAWNEWSCMNCGIGTGTTRVACVSP 1395

RESULT 10
A39804
thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A39804
R:LaBelle, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A:Molecule type: mRNA
A:Residues: 1-1178 <LAB>
A:Cross-references: UNIPROT:P35440; GB:M60853; NID:G212763; PIDN:AAA51437.1; PID:G212764
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von
F:325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F:386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F:499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F:658-697/Domain: EGF homology <EGF>

Query Match 37.9%; Score 466; DB 1; Length 1178;
Best Local Similarity 43.9%; Pred. No. 1.8e-27;
Matches 87; Conservative 19; Mismatches 76; Indels 16; Gaps 3;

QY 5 GFSQSAWRACSVTCGKGIKSRSLCNPPLPANGKPCQGSLEMRNCQKPCP-----VD 60
DB 389 GWSFSDWTCKVTCSGTQMRGSCDVTRSA-----CTGPHIQTRMCSFKKCDHRIQD 443

QY 61 GSWSESLWEECTRCGRNQTRTRTCNNPVSQHGRCPEGNAVEIIMCNRPCPVHGAW 120
DB 444 GWSHSPWSSCSVTCGVGNITRILCNSPIQMGKXKCVNGRETKCEKAPCPVNGQW 503

QY 121 SAWQPGWTCSECGKGTQTRARLCNNPPPAFGSGYCDGAETOMQVNCERNICPHG----- 175
DB 504 GWSFWSACTVTCGGGIRSRSLCNSPEPYGKPCVGTQKQHDNCRKRDCEPIDGCLSNP 563

QY 176 --KWATWASWACSVCSCG 191
DB 564 CFFGAECSYPPDGWSWSCG 581

RESULT 11
TSHP2
thrombospondin 2 precursor - human
C:Species: Homo sapiens (man)
C>Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: A47379; A42173
R:LaBelle, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
A:Reference number: A47379; MUID:94010892; PMID:8406456
A:Accession: A47379
A:Molecule type: mRNA
A:Residues: 1-1172 <LAB>
A:Cross-references: UNIPROT:P35442; GB:IL12350; NID:G307505; PIDN:AAA03703.1; PID:G307506
R:LaBelle, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression o
A:Reference number: A42173; MUID:92217961; PMID:1559694
A:Accession: A42173
A:Molecule type: mRNA
A:Residues: 560-1172 <LA2>
```

```
A:Cross-references: GB:M81339
A:Experimental source: fibroblast
A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIIP:95096)
C:Genetics:
A:Gene: GDB:THBS2; TSP2
A:Cross-references: GDB:128789; OMIM:188061
A:Map position: 6q27-6q27
C:Complex: homotrimer, disulfide linked
C:Function:
A:Description: participates in cell migration and adhesion, and in platelet aggregation
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF>
F:652-691/Domain: EGF homology <EGF>
F:928-930/Region: cell attachment (R-G-D) motif
F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predic
F:167-226/Disulfide bonds: #status predicted
F:266-270/Disulfide bonds: interchain #status predicted
F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 36.8%; Score 453; DB 1; Length 1172;
Best Local Similarity 40.4%; Pred. No. 1.6e-26;
Matches 80; Conservative 26; Mismatches 76; Indels 16; Gaps 3;

QY 5 GFSQSAWRACSVTCGKGIKSRSLCNPPLPANGKPCQGSLEMRNCQKPCPV---D 60
DB 383 GWSFWSACTVTCSGTQMRGSCDVTRSA-----SVTCLGPSICTRACSLKCDTRIQRD 437

QY 61 GSWSESLWEECTRCGRNQTRTRTCNNPVSQHGRCPEGNAVEIIMCNRPCPVHGAW 120
DB 438 GWSHSPWSSCSVTCGVGNITRILCNSPVQMGKXKCVNGRETKCEKAPCPIDGRW 497

QY 121 SAWQPGWTCSECGKGTQTRARLCNNPPPAFGSGYCDGAETOMQVNCERNICPHG----- 175
DB 498 SPWSPWACTVTCAGGIRSRTRVCNSPEPYGKPCVGTQKQHDNCRKRDCEPIDGCLSNP 557

QY 176 --KWATWASWACSVCSCG 191
DB 558 CFFGAECSYPPDGWSWSCG 575

RESULT 12
A42587
thrombospondin 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A42587; A39851
R:LaBerty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: A42587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1172 <LAH>
A:Cross-references: UNIPROT:Q03350; GB:L07803; GB:M87275; NID:G340421; PIDN:AAA53064.1;
A:Note: sequence extracted from NCBI backbone (NCBIIP:81502)
R:Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A:Reference number: A39851; MUID:91302287; PMID:1712771
A:Accession: A39851
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-873 <BOR>
A:Cross-references: GB:M64866; NID:G201994; PIDN:AAA40432.1; PID:G201995
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; vo
```

A;Residues: 1-60, 'X', 62-413, 'D', 415-452, 'XX', 455-469 <WE2>

Best Local Similarity 35.9%; pred. No. 1.1e-25;

	Matches	85;	Conservative	22;	Mismatches	80;	Indels	50;	Gaps	5;
QY	1	QVHGFSQSAWRACSVTCGKGIQKRSRLCNQP	-----	LPANG	38					
Db	657	RVGNESWTWENHCSVSGRGSQARYKCLSPHRTLAFDCPGENKVTNELRITFFKARS	716							
QY	39	GKPC-----	QSDLEMRNCQKPCPVDGSWSWSLWEECTSCGRGNQTRRT	86						
Db	717	YIMCSVRCKIKRNTISEKNIEVRSNDGPNCAIGWGTWGWSTCSTSCGPGTLVRQT	776							
QY	87	CNNPSVQHGGRCPEGNAVEIIMCNIRPCPVHGAWSAQWPWGTCSES CGKGTQTRARLCNN	146							
Db	777	CNR-----	EPCDGAHRRSCNVATCQNDGINSLNWNSDCSRVCGKGLRSRSRSC--	827						
QY	147	PPPAFGSYCDGAETQMVQCNERNCPH--GKWATWASWSACSVSCGGGARQTRGC	201							
Db	828	-----	FGSCMGASSEQQCFNEQACASSANDWGTWSGWSQCSVSCGAGVKRTRTC	879						

RESULT 15

T25061

hypothetical protein T21B6.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T25061

R:Cottage, A.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19975

A:Accession: T25061

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-788 <WIL>

A:Cross-references: UNIPROT:Q22631; EMBL:Z68011; PIDN:CAA92014.1; GSPDB:GN00028; CESP:T2

A:Experimental source: clone T21B6

C:Genetics:

A:Gene: CESP:T21B6.3

A:Map position: X

A:Introns: 20/1; 47/1; 76/1; 152/1; 735/2; 754/2

Query Match 33.9%; Score 417.5; DB 2; Length 788;

Best Local Similarity 39.0%; Pred. No. 5.3e-24;

	Matches	78;	Conservative	20;	Mismatches	79;	Indels	23;	Gaps	6;
QY	2	VHGGFSQSAWRACSVTCGKGIQKRSRLCNQP	LPANGKPKCGSDLEMRNCQKPCPVDG	61						
Db	454	VSGVHWDSDWSTCTCGDGAKSRRECST-----	NNCQADYETEPFNLGFCQ---	503						
QY	62	SWSESLWEECTSCGRGNQTRTCNNFNVQHGGRPCEGNAVEIIMCNIRPCPVHGAWS	121							
Db	504	TWSEWCWSTCSASCSGSGQRERTRFCH-----	LGITNRCGKDYSESQCSAGPCP---	555						
QY	122	AWQPMGTCSGCGKGTQTRARLCNNPPPAFGGSYCDGAETQMVQCNERNCPHGWATWA	181							
Db	556	QWEDWQCQSVTCGGVAVRQRTCLG--GVFGDHLCOGPKTEQACDGGFCSL---	WSPWQ	610						
QY	182	SWSACSVSCGGGARQTRGC	201							
Db	611	EWSTGASCGSGMKRRQRVC	630							

Search completed: November 17, 2004, 15:56:20

Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2004, 15:30:24 ; Search time 194 Seconds

(without alignments)
616.897 Million cell updates/sec

Title: US-10-019-065A-1

Perfect score: 1231

Sequence: 1 QVHGFGFSQSAWRACSVTCG.....SCGGGARQTRGCDPPVQY 208

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231	100.0	5636	2 Q96RW7	Q96RW7 homo sapien
2	1230	99.9	2673	2 Q96SC3	Q96SC3 homo sapien
3	533.5	43.3	1077	1 S5A_MOUSE	Q62217 mus musculus
4	530.5	43.1	1582	2 Q8CGM0	Q8CGM0 mus musculus
5	528.5	42.9	1388	2 Q7QKD0	Q7QKD0 anopheles g
6	522.5	42.4	1074	1 S5A_HUMAN	Q13591 homo sapien
7	519.5	42.2	1584	1 BAI1_HUMAN	Q14514 homo sapien
8	511.5	41.6	1093	1 S5B_HUMAN	Q9p283 homo sapien
9	511.5	41.6	1151	2 Q6DD89	Q6dd89 homo sapien
10	507.5	41.2	478	2 Q8BVE5	Q8bve5 m mus muscu
11	507.5	41.2	632	2 Q62PQ8	Q62pq8 mus musculus
12	507.5	41.2	632	2 BAC98172	Bac98172 mus muscu
13	507.5	41.2	1093	1 S5B_MOUSE	Q60519 mus musculus
14	506.5	41.1	1088	2 Q6PCK8	Q6pck8 xenopus lae
15	506.5	41.1	1088	2 AAH59288	AAH59288 xenopus l
16	504.5	41.0	612	2 Q6ZQ96	Q6zq96 mus musculus
17	504.5	41.0	612	2 BAC97972	Bac97972 mus muscu
18	504.5	41.0	1522	1 BAI3_MOUSE	Q80zf8 mus musculus
19	503.5	40.9	1522	1 BAI3_HUMAN	O60242 homo sapien
20	500.5	40.7	447	2 Q6DCS2	Q6dcs2 xenopus lae
21	496.5	40.3	1549	2 Q6PQ80	Q6pq80 mus musculus
22	496.5	40.3	1549	2 AAH56926	AAH56926 mus muscu
23	492.5	40.0	1170	1 TSPI_BOVIN	Q28178 bos taurus
24	492.5	40.0	1170	1 TSPI_MOUSE	P35441 mus musculus
25	492.5	40.0	1171	2 Q80YQ1	Q80yq1 mus musculus
26	492.5	40.0	1171	2 Q8CGB2	Q8cgb2 mus musculus
27	492.5	40.0	1560	2 Q8CGM1	Q8cgm1 mus musculus
28	490	39.8	1092	2 Q6UY12	Q6uy12 homo sapien
29	490	39.8	1092	2 AAQ88491	AAQ88491 homo sapi
30	489.5	39.8	1572	1 BAI2_HUMAN	O60241 homo sapien
31	489.5	39.8	1573	2 Q8NGW8	Q8ngw8 homo sapien

ALIGNMENTS

RESULT 1

ID	Q96RW7	PRELIMINARY;	PRT; 5636 AA.
AC	Q96RW7;		
DT	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)		
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)		
DE	Hemicentin.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Trent J.;		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF156100; AAK68690.1; -.		
DR	HSSP; P07996; 1LSL.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0005509; F:calcium ion binding; IEA.		
DR	GO; GO:0005021; F:vascular endothelial growth factor receptor. .; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	InterPro; IPR00152; Asx_hydroxyl_S.		
DR	InterPro; IPR000875; Cecropin.		
DR	InterPro; IPR000742; EGF_2.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	InterPro; IPR006209; EGF_like.		
DR	InterPro; IPR006605; G2F.		
DR	InterPro; IPR009030; Grow_fac_recept.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003598; Ig_C2.		
DR	InterPro; IPR000169; Pept_cys_acsite.		
DR	InterPro; IPR000884; TSPI.		
DR	InterPro; IPR009134; VEGFR.		
DR	InterPro; IPR002035; VWF_A.		
DR	Pfam; PF07645; EGF_CA; 8.		
DR	Pfam; PF07474; G2F; 1.		
DR	Pfam; PF00047; Ig; 44.		
DR	Pfam; PF00090; TSP_1; 6.		
DR	PRINTS; PR01832; VEGFRECEPTOR.		
DR	SMART; SM00179; EGF_CA; 7.		
DR	SMART; SM00408; IG2; 43.		
DR	SMART; SM00209; TSPI; 6.		
DR	SMART; SM00327; VWA; 1.		
DR	PROSITE; PS00010; ASX_HYDROXYL; 5.		
DR	PROSITE; PS00268; CECROPIN; UNKNOWN_1.		
DR	PROSITE; PS01186; EGF_2; 3.		
DR	PROSITE; PS00266; EGF_3; 5.		
DR	PROSITE; PS01187; EGF_CA; 8.		
DR	PROSITE; PS06835; IG_LIKE; 44.		
DR	PROSITE; PS06639; THIOL_PROTEASE_HIS; UNKNOWN_1.		
DR	PROSITE; PS00092; TSPI; 6.		
KW	EGF-like domain.		
SC	SEQUENCE 5636 AA; 613673 MW; F000B319CED7B52C CRC64;		

```

Query Match      100.0%; Score 1231; DB 2; Length 5636;
Best Local Similarity 100.0%; Pred. No. 3.2e-89;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQMSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKCGSDLEMRNCQKPCPVD 60
Db 4528 QVHGFSQMSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKCGSDLEMRNCQKPCPVD 4587
QY 61 GSWSEWSLWEECTRS CGRGNGQTRTTCNNPVSQHGGRPCGNAVEIIMCNRCPVHGAW 120
Db 4588 GSWSEWSLWEECTRS CGRGNGQTRTTCNNPVSQHGGRPCGNAVEIIMCNRCPVHGAW 4647
QY 121 SAWOPWGTCTSES CGKGTQTRARLNCNNPPAFGGSYCDGAEQVQVNCNRCPIHGKWATW 180
Db 4648 SAWOPWGTCTSES CGKGTQTRARLNCNNPPAFGGSYCDGAEQVQVNCNRCPIHGKWATW 4707
QY 181 ASWSACSVSCGGGARQRTGCSDPVPOY 208
Db 4708 ASWSACSVSCGGGARQRTGCSDPVPOY 4735

RESULT 2
Q96SC3 PRELIMINARY; PRT; 2673 AA.
AC Q96SC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fibulin-6 (Fragment).
GN Name=FIBL-6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Melanoma;
RA Kostka G., Timpi R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306906; C37830.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR000152; Asx hydroxyl_S.
DR InterPro; IPR000875; Cectopin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006605; G2F.
DR InterPro; IPR009017; GFF like.
DR InterPro; IPR009030; Grow fac recept.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR009138; Neural_cell_adh.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF07645; EGF_CA; 8.
DR Pfam; PF07474; G2F; 1.
DR Pfam; PF00047; Ig; 17.
DR Pfam; PF00090; TSP_1; 6.
DR PRINTS; PR01838; NCAMFAMILY.
DR PRINTS; PR01705; TSPIREPEAT.
DR SMART; SM00179; EGF_CA; 7.
DR SMART; SM00408; IGC2; 17.
DR SMART; SM00209; TSPI; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS00268; CECTOPIN; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS00026; EGF_3; 5.
DR PROSITE; PS01187; EGF_CA; 8.
DR PROSITE; PS00835; IGLIKE; 17.
DR PROSITE; PS00092; TSPI; 6.

```

```

KW EGF-like domain.
FT NON TER 1
SQ SEQUENCE 2673 AA; 291017 MW; BEABC30B8340E272 CRC64;

Query Match      99.9%; Score 1230; DB 2; Length 2673;
Best Local Similarity 99.5%; Pred. No. 1.9e-89;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVHGFSQMSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKCGSDLEMRNCQKPCPVD 60
Db 1565 QVHGFSQMSAWRACSVTCGKGIQKRSRLCNQPLPANGKPKCGSDLEMRNCQKPCPVD 1624
QY 61 GSWSEWSLWEECTRS CGRGNGQTRTTCNNPVSQHGGRPCGNAVEIIMCNRCPVHGAW 120
Db 1625 GSWSEWSLWEECTRS CGRGNGQTRTTCNNPVSQHGGRPCGNAVEIIMCNRCPVHGAW 1684
QY 121 SAWOPWGTCTSES CGKGTQTRARLNCNNPPAFGGSYCDGAEQVQVNCNRCPIHGKWATW 180
Db 1685 SAWOPWGTCTSES CGKGTQTRARLNCNNPPAFGGSYCDGAEQVQVNCNRCPIHGKWATW 1744
QY 181 ASWSACSVSCGGGARQRTGCSDPVPOY 208
Db 1745 ASWSACSVSCGGGARQRTGCSDPVPOY 1772

RESULT 3
SMSA_MOUSE
ID SMSA_MOUSE STANDARD; PRT; 1077 AA.
AC Q62217;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN Name=Sema5a; Synonyms=Semaf, SemF;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NMRL;
RA MEDLINE=96414430; PubMed=8817451;
RA Adams R.H., Betz H., Puschel A.W.;
RT "A novel class of murine semaphorins with homology to thrombospondin
RT is differentially expressed during early embryogenesis.";
RL Mech. Dev. 57:33-45(1996).
CC -!- FUNCTION: May act as positive axonal guidance cues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: In adult, detected in liver, brain, kidney,
CC heart, lung and spleen.
CC -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
CC adult tissues. Its abundance decreases from E10 to birth.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X97817; CAA66397.1; -.
CC HSSP; P07996; 1LSL.
CC MGD; MGI:107556; Sema5a.
CC GO; GO:0016021; C:integral to membrane; IDA.
CC GO; GO:0008046; P:axon guidance; IMP.
CC GO; GO:0007411; P:axon guidance; IMF.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR002165; Plexin_repeat.
CC InterPro; IPR001627; Sema.

```


DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP_1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; TSP_1; 5.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP1; 6.
DR PROSITE; PS51004; Sema; 1.
DR PROSITE; PS50092; TSP1; 6.
KW Developmental protein; Glycoprotein; Multigene family; Neurogenesis;
KW Repeat; Signal; Transmembrane.
FT SIGNAL; 1 21
FT CHAIN 22 1077
FT DOMAIN 22 971
FT CHAIN 22 971
FT DOMAIN 22 971
FT TRANSMEM 992 992
FT DOMAIN 993 1077
FT DOMAIN 35 484
FT DOMAIN 540 593
FT DOMAIN 595 651
FT DOMAIN 653 702
FT DOMAIN 707 765
FT DOMAIN 784 839
FT DOMAIN 841 896
FT DOMAIN 897 944
FT CARBOHYD 147 147
FT CARBOHYD 168 168
FT CARBOHYD 227 227
FT CARBOHYD 277 277
FT CARBOHYD 323 323
FT CARBOHYD 367 367
FT CARBOHYD 536 536
FT CARBOHYD 591 591
FT CARBOHYD 717 717
FT CARBOHYD 933 933
SQ SEQUENCE 1077 AA; 120826 MW; EDABDDDA42789FF CRC64;

Query Match 43.3%; Score 533.5; DB 1; Length 1077;
Best Local Similarity 34.1%; Pred. No. 2.7e-34;
Matches 98; Conservative 23; Mismatches 79; Indels 87; Gaps 3;

QY 3 HGFQSWSAWRACVTCGKGIQKRSRLCNQPLPANGKFCQGSDLERNKNC-QNKPCPVDG 61
Db 595 NGGWTPTWSVSPCTTCGIGFQVRQSCGNFTPRHGGVCGVGNREERYCNEHLLCPHV 654
QY 62 SWSWSLWEECTSCRGNGQTRTRTCNNFVSQHGGRPCGNAVEIIMCNIRPCP----- 115
Db 655 FWTGWFWESECTACCGGIGQARRTCN-----GPDCAGSNVEYHPCNTNACPELKXTT 708
QY 116 ----- 115
Db 709 PWTPTVNIISDNGHYEQRYFTCKARLPDPNLLVGRQRIEMRYCSDGTSGCSTDGL 768
QY 116 -----VHGAWSAWQFWGTCSESCGKGTQTRARLNNPPAPFGGSCDGAET 161
Db 769 SGDFLRAGRYSAHTVANGAWSAWTSWQSCDRCGRIRNRKRVNPNPEKFGGMPCLGPSL 828
QY 162 QMVQCNERNCPHKGWATWASWACSVSCGGARQTRGSDRVPQY 208
Db 829 EFQECNLPFCVDGVNSWNSWKSATCGGHHYMTSCSNAPAY 875

RESULT 4
Q8CGMO PRELIMINARY; PRT; 1582 AA.
ID Q8CGMO
AC Q8CGMO
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Brain-specific angiogenesis inhibitor 1.
GN Name=Bail;
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RX MEDLINE=21143048; PubMed=11245925;
RA Koh J.T., Lee Z.H., Ahn K.Y., Kim J.K., Bae C.S., Kim H.H., Kee H.J.,
RA Kim K.K.;
RT "Characterization of mouse brain-specific angiogenesis inhibitor 1
RT (Bail) and phytanoyl-CoA alpha-hydroxylase-associated protein 1, a
RT novel Bail-binding protein.";
RL Brain Res. Mol. Brain Res. 87:223-237(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=ICR; TISSUE=Brain;
RC Koh J.T., Kee H.J., Kim K.K.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY168408; AAN86966.1; -;
DR HSSP; P07996; 1LSL.
DR MGD; MGI:1933736; Bail.
DR GO; GO:0005515; P:protein binding; IPI.
DR InterPro; IPR008077; Angio_inhib.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP_1.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00090; TSP_1; 5.
DR PRINTS; PR01694; BAIPRECURSOR.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50221; GPS; 1.
DR PROSITE; PS50227; G_PROTEIN_RECP_F2_3; 1.
DR PROSITE; PS50261; G_PROTEIN_RECP_F2_4; 1.
DR PROSITE; PS50092; TSP1; 4.
SQ SEQUENCE 1592 AA; 173111 MW; FACE62C009F02F8E CRC64;

Query Match 43.1%; Score 530.5; DB 2; Length 1592;
Best Local Similarity 46.8%; Pred. No. 6.7e-34;
Matches 94; Conservative 25; Mismatches 71; Indels 11; Gaps 5;

QY 8 QWSAWRACVTCGKGIQKRSRLCNQPLPANGKFCQGSDLERNKNCQKP-CPVDGWSWSEW 66
Db 359 EWSPWSVCSSTCGSGWQTRTRFC---VSSYSTQCSGPLREQLRNLNNSAVCPVHGAWDEW 415
QY 67 SLWEECTSCRGNGQTRTRTCNNFVSQHGGRPCGNAVEIIMCNIRPCP---VHGAWSAW 123
Db 416 SPWSLCSSTCGRFPDRTRTCRPP--QFGNPGCEGPEKQTKFCNLALCPGRAVDGNWNEW 473
QY 124 QPWGTCSESCGKGTQTRARLNNPPAPFGGSCDGAETQMVQCNERNCPHKGWATWASW 183
Db 474 SWSWSTCSASCSCGGRQTRTCNG--PSYGGAECCQGHVWETRDCLQCCPVDGKQWASW 531
QY 184 SACSVCSCGGARQTRGSDP 204
Db 532 GCSVTCGGSGQRREVCSP 552

RESULT 5
Q7QKDO PRELIMINARY; PRT; 1388 AA.
ID Q7QKDO
AC Q7QKDO
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AgCP14698 (Fragment).

GN Name=agCG49943; ORFNames=ENSANGG00000019365;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
RX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008799; EAA03790.1; --
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00047; Ig; 6.
DR Pfam; PF00090; TSP 1; 4.
DR PRINTS; PR01705; TSP1REPEAT.
DR PROSITE; PS50835; IG LIKE; 9.
DR PROSITE; PS50092; TSP1; 4.
FT NON TER 1388 1388
SQ SEQUENCE 1388 AA; 153015 MW; 3A054C8EF1900324 CRC64;
Query Match 42.9%; Score 528.5; DB 2; Length 1388;
Best Local Similarity 42.2%; Pred. No. 8.6e-34;
Matches 89; Conservative 36; Mismatches 77; Indels 9; Gaps 3;
Qy 6 FQSWASVAVRACSVTCGKGIQKRSRLC--NQPLPANGGK-PQGSDELVRMCNQKPCPDVGS 62
Db 1072 WSPGAMSPCSATCGSGTQFRSRICLLVNGSPAHGERFNCVGVNELKACELLPCPDVNG 1131
Qy 63 WSESLVEETRSC-----GRGNQTRTCNNPSVQHGRPCGKNAVILMCNIRPCV 116
Db 1132 WGEWTGWSNLSLSCVSEFSGVRSIRKSRACDPAPSLGKQCVGAYEEPCHKYICPI 1191
Qy 117 HGASVAVRACSVTCGSGKGTQTRALCNPPPAFGSGYCDGAEQVQVNCNRCPIHGK 176
Db 1192 DGGWTAWSWTGCSEPCGFCGSLRMSRCSNPVPRHGLGLCDGAESEVKACKVQECHVDG 1251
Qy 177 WATWASVAVRACSVCGGAGARTRGCSDPVQ 207
Db 1252 WSEWRSPCNKSGGKIKRRRYCNNEPK 1282
RESULT 6
SM5A_HUMAN STANDARD; PRT; 1074 AA.
AC Q13591; Q60408;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
GN Name=SEMA5A; Synonyms=SEMAF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98125554; PubMed=9464278;
RA Simmons A.D.; Fuschel A.W.; McPherson J.D.; Overhauser J.; Lovett M.;
RT "Molecular cloning and mapping of human semaphorin F from the Cri-du-
RT chat candidate interval."
RL Biochem. Biophys. Res. Commun. 242:685-691(1998).
RN [2]
RC SEQUENCE OF 1-494 FROM N.A.
RA Kalicki J.; Harmon G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May act as positive axonal guidance cues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the semaphorin family.

CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; U52840; AAC09473.1; --
DR EMBL; AC004615; AAC14668.1; --
DR FIR; JC5928; JC5928.
DR HSP; P07996; ILSL.
DR HSP; HGNC:10736; SEMA5A.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007393; P:neurogenesis; TAS.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; TSP 1; 6.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP1; 6.
DR PROSITE; PS51004; SEMA; 1.
DR PROSITE; PS50092; TSP1; 6.
KW Developmental protein; Glycoprotein; Multigene family; Neurogenesis;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1074 Semaphorin 5A.
FT DOMAIN 23 968 Extracellular (Potential).
FT TRANSMEM 969 989 Potential.
FT DOMAIN 990 1074 Cytoplasmic (Potential).
FT DOMAIN 35 484 Sema.
FT DOMAIN 540 593 TSP type-1 1.
FT DOMAIN 595 651 TSP type-1 2.
FT DOMAIN 653 702 TSP type-1 3.
FT DOMAIN 707 765 TSP type-1 4.
FT DOMAIN 784 839 TSP type-1 5.
FT DOMAIN 841 896 TSP type-1 6.
FT DOMAIN 897 944 TSP type-1 7.
FT CARBOHYD 142 142 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 168 168 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 227 227 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 277 277 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 323 323 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 367 367 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 437 437 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 536 536 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 591 591 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 717 717 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 933 933 N-linked (GlcNAc...) (Potential).
FT CONFLICT 56 56 A -> V (in Ref. 2).
FT CONFLICT 149 149 A -> T (in Ref. 2).
FT CONFLICT 382 382 V -> M (in Ref. 2).
FT CONFLICT 494 494 S -> R (in Ref. 2).
SQ SEQUENCE 1074 AA; 120570 MW; EE3DB763CBE29407 CRC64;
Query Match 42.4%; Score 522.5; DB 1; Length 1074;
Best Local Similarity 33.1%; Pred. No. 2e-33;
Matches 95; Conservative 24; Mismatches 81; Indels 87; Gaps 3;
Qy 3 HGFSSQSWAVRACSVTCGKGIQKRSRLCNPQLPANGGKPCQGSDELVRMC-QNKPCPDV 61
Db 595 NGWTPWTSWSPCSTTCGIGFQVRQSCSNPTPRHGGRVCGVQGRREYCNHLLCPPHM 654

QY 62 SWSWLSWECTSCRGNGQTRTCNNPNSVOHGRPCBGNVAIVEIMCNIIRCP----- 115
Db 655 FWTGWFWRCTACQCGGIGQARRICEN-----GPDACGCVVEYQSCNTNCPBLKXTT 708
QY 116 ----- 115
Db 709 PWTPTVNVISDNGDHYEQFRYTKCARLADPNLLEVGQRLEMYCSDGTSGCSTDGL 768
QY 116 -----VHGANSANQPMWCTSCSGCKGFTQTRARICNNPPPAFGSGYCDGAET 161
Db 769 SGDFLAGRYSARTVNGASAVTSQCSRDGSRGIRNKRVCNNPEPKYGMPCIGPGL 828
QY 162 OMQVCNERNCPHKGKATWASACSVCSCGGARORTGCSDPVPOY 208
Db 829 EYQECNTLPCPDGVWSCWSPWIKSATCGGHYMRTRSCSNPAPAY 875
RESULT 7
BAIL_HUMAN STANDARD; PRT; 1584 AA.
ID AC 014574;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Brain-specific angiogenesis inhibitor 1 precursor.
GN Name=Bail;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Petal brain;
RX MEDLINE=98054121; PubMed=9393972;
RA Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,
RA Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;
RT "A novel brain-specific p53-target gene, BAIL, containing
thrombospondin type 1 repeats inhibits experimental angiogenesis.";
RL Oncogene 15:2145-2150(1997).
RN [2]
RP INTERACTION WITH BAP1.
RX MEDLINE=98321173; PubMed=9647739;
RA Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,
RA Tokino T.;
RT domain-containing protein that interacts with BAIL.";
RL Biochem. Biophys. Res. Commun. 247:597-604(1998).
CC -!- FUNCTION: Likely to be a potent inhibitor of angiogenesis in brain
and may play a significant role as a mediator of the p53 signal in
suppression of glioblastoma. May function in cell adhesion and
signal transduction in the brain.
CC -!- SUBUNIT: Interacts with BAP1 and PHYHIP.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Likely to be
concentrated at cell-cell adhesion sites.
CC -!- TISSUE SPECIFICITY: Specifically expressed in brain. Reduced or no
expression is observed in some glioblastoma cell lines and cancer
tissues.
CC -!- INDUCTION: By p53.
CC -!- DOMAIN: The TSP1 repeats inhibit in vivo angiogenesis in rat
cornea induced by BFGF.
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC -!- LN-TM7 subfamily.
CC -!- SIMILARITY: Contains 1 GPS domain.
CC -!- SIMILARITY: Contains 5 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----

EMBL; AB005297; BAA23647.1; -.
PIR; T00026; T00026.
HSSP; P07996; ILSL.
Genew; HGNC:943; BAIL.
MIM; 602682; -.
GO; GO:0005687; C:integral to plasma membrane; TAS.
GO; GO:0005911; C:intercellular junction; TAS.
GO; GO:0005515; F:protein binding; TAS.
GO; GO:0007409; P:axogenesis; TAS.
GO; GO:0007155; P:cell adhesion; TAS.
GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
GO; GO:0007422; P:peripheral nervous system development; TAS.
GO; GO:0007155; P:signal transduction; TAS.
InterPro; IPR008077; Angio_inhib.
InterPro; IPR008032; GPCR_secretin.
InterPro; IPR001879; Hormn_receptor.
InterPro; IPR00203; PKD_cys_rich.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP_1.
Pfam; PF00002; 7tm_2; 1.
Pfam; PF01825; GPS; 1.
Pfam; PF02793; HRM; 1.
Pfam; PF00090; TSP_1; 5.
PRINTS; PR01694; BAIPRECURSOR.
PRINTS; PR00249; GPCRSECRETIN.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00303; GPS; 1.
SMART; SM00008; Hormr; 1.
SMART; SM00209; TSP1; 5.
PROSITE; PS0221; GPS; 1.
PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE NEG.
PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE NEG.
PROSITE; PS0227; G_PROTEIN_RECEP_F2_3; 1.
PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
PROSITE; PS0092; TSP1; 5.
Cell adhesion; G-protein coupled receptor; Glycoprotein; Repeat;
Signal; Transmembrane.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 1584 Brain-specific angiogenesis inhibitor 1.
FT DOMAIN 31 948 Extracellular (Potential).
FT TRANSMEM 949 969 1 (Potential).
FT DOMAIN 970 980 Cytoplasmic (Potential).
FT TRANSMEM 981 1001 2 (Potential).
FT DOMAIN 1002 1008 Extracellular (Potential).
FT TRANSMEM 1009 1029 3 (Potential).
FT DOMAIN 1030 1052 Cytoplasmic (Potential).
FT TRANSMEM 1053 1073 4 (Potential).
FT DOMAIN 1074 1093 Extracellular (Potential).
FT TRANSMEM 1094 1114 5 (Potential).
FT DOMAIN 1115 1136 Cytoplasmic (Potential).
FT TRANSMEM 1137 1157 6 (Potential).
FT DOMAIN 1158 1166 Extracellular (Potential).
FT TRANSMEM 1167 1187 7 (Potential).
FT DOMAIN 1188 1584 Cytoplasmic (Potential).
FT TRANSMEM 1584 1584 TSP type-1 1.
FT DOMAIN 354 407 TSP type-1 2.
FT TRANSMEM 409 462 TSP type-1 3.
FT DOMAIN 467 520 TSP type-1 4.
FT TRANSMEM 522 575 TSP type-1 5.
FT DOMAIN 881 938 GPS.
FT TRANSMEM 1411 1422 Poly-Pro.
FT DOMAIN 1425 1430 Poly-Pro.
FT SITE 231 233 Cell attachment site (Potential).
FT DOMAIN 1365 1584 Necessary for interaction with BAP1.
FT TRANSMEM 1581 1584 Indispensable for interaction with BAP1.
FT TRANSMEM 64 64 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 401 401 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 607 607 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 692 692 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 844 844 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 877 877 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 881 881 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1584 AA; 173531 MW; DEAF28C77874513 CRC64;

Query Match 42.2%; Score 519.5; DB 1; Length 1584;
 Best Local Similarity 45.0%; Pred. No. 5.1e-33;
 Matches 92; Conservative 25; Mismatches 73; Indels 11; Gaps 5;

QY 8 QWAGWACVTCGKGTQKSRCLNQLPANGGKPCGSDLEMCNKP-CPVDGWSW 66
 Db 359 EWGFWSCSTGEGWQTRFC---VSSYSTQCSGRLPRLNNSAVCPVHGAWDEW 415
 QY 67 SLMEETRSOGRGNQTRTTCNNPSVQHGRPCGEGNAVEIIMCNIRPCP---VHGAWSAW 123
 Db 416 SPNSLCSSTCGRFRDTRTCRPP--QFGNPGCEGPEKQTKFCNIALCPGRAVDGNWNEW 473
 QY 124 QPVGTCSECKGTOTRFLCNPPAFGSGYCDGAEQVQVNCNERNCPHKGWATWASW 183
 Db 474 SWSACSASCSQGRQRTRECNG--PSYGBACQGHVETRDCLQQCPVDGKQWQWASW 531
 QY 184 SACSVCGGGARQRTGCSDP 204
 Db 532 GSCSVTCGAGSQRRVCSGP 552

RESULT 8
 SWSB HUMAN STANDARD; PRT; 1093 AA.
 ID _SWSB HUMAN STANDARD; PRT; 1093 AA.
 AC Q9P283;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Semaphorin 5B precursor.
 GN Name=SEMA5B; Synonyms=KIRA1445;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20277482; PubMed=10819331;
 RA Nagase T, Kikuno R, Ishikawa K-I, Hirose M, Ohara O;
 RT "Prediction of the coding sequences of unidentified human genes. XVII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 7:143-150(2000).
 CC -!- FUNCTION: May act as positive axonal guidance cues (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: Belongs to the semaphorin family.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB040878; BAA95969.1; ALT_INIT.
 CC GenBank: U000000000.1; SEMA5B.
 CC InterPro: IPR003659; Flexin-like.
 CC InterPro: IPR002165; Flexin_repeat.
 CC InterPro: IPR001627; Sema.
 CC InterPro: IPR000884; TSP1.
 CC InterPro: IPR008085; TSP1.
 CC Pfam: PF01437; PSI; 1.
 CC Pfam: PF01403; Sema; 1.
 CC Pfam: PF00090; TSP1; 5.
 CC PRINTS: PR01705; TSP1REPEAT.
 CC SMART: SM00423; PSI; 1.
 CC SMART: SM00630; Sema; 1.
 CC SMART: SM00209; TSP1; 5.

DR PROSITE; PSS1004; SEMA; 1.
 DR PROSITE; PSS0092; TSP1; 5.
 KW Developmental protein; Glycoprotein; Multigene family; Neurogenesis;
 KW Repeat; signal; Transmembrane.
 FT SIGNAL 1 26 Potential.
 FT CHAIN 27 1093 Semaphorin 5B.
 FT DOMAIN 20 978 Extracellular (Potential).
 FT TRANSMEM 979 999 Potential.
 FT DOMAIN 1000 1093 Cytoplasmic (Potential).
 FT DOVAIN 45 495 Sema.
 FT DOVAIN 551 605 TSP type-1 1.
 FT DOVAIN 606 662 TSP type-1 2.
 FT DOVAIN 664 713 TSP type-1 3.
 FT DOVAIN 721 776 TSP type-1 4.
 FT DOVAIN 795 850 TSP type-1 5.
 FT DOVAIN 852 907 TSP type-1 6.
 FT DOVAIN 908 952 TSP type-1 7.
 FT CARBOHYD 59 59 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 95 95 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 157 157 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 178 178 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 287 287 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 333 333 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 378 378 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 532 532 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 539 539 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 547 547 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 602 602 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 728 728 N-linked (GlcNAc... (Potential).
 FT CARBOHYD 944 944 N-linked (GlcNAc... (Potential).
 SQ SEQUENCE 1093 AA; 119866 MW; F1DFBF97CEAF0EF CRC64;
 Query Match 41.6%; Score 511.5; DB 1; Length 1093;
 Best Local Similarity 33.0%; Pred. No. 1.6e-32;
 Matches 94; Conservative 22; Mismatches 82; Indels 87; Gaps 3;

QY 3 HGFQSWACVTCGKGTQKSRCLNQLPANGGKPCGSDLEMCNKP-CPVDG 61
 Db 606 NGAWTPWSSWALCSTGCGIQTQVQRSCSNPAPRHGRI CVGSRERFCNENTPCVPI 665
 QY 62 SWSWSLWEECTRCGRGNQTRTTCNNPSVQHGRPCGEGNAVEIIMCNIRPCP----- 115
 Db 666 FWAGWSKRSKSSNCGGMSRRACEN-----GNSCLGCVFEKTCNPEGCEVRNT 719
 QY 116 ----- 115
 Db 720 PWTPLPVPVNTGGARQORFRETCTAPLADPHGLQFGRERTTCTCPADGSGCDTAL 779
 QY 116 -----VHGAWGAWQFWGTCSECGKGTQTRARLNNPPAPFGGYCDGAET 161
 Db 780 VEDLLRGSTSPHTVSGWAWGFWSSCSRDCELGFPVRKCTCTNPEPRNGGLPCVGDA 839
 QY 162 QMQVCNERNCPHKGWATWASWSACSVCSSGCGARQRTGCDPVP 206
 Db 840 EYQDCNQCACPVGAWSCWTSWSPSCSACGCGHYQRTSCTSPAP 884

RESULT 9
 Q6DD89
 ID Q6DD89 PRELIMINARY; PRT; 1151 AA.
 AC Q6DD89;
 DT 01-OCT-2004 (Tremblrel. 28, Created)
 DT 01-OCT-2004 (Tremblrel. 28, Last sequence update)
 DT 01-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE SEMA5B protein.
 GN Name=SEMA5B;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;

MEDLINE=9279253; PubMed=10349636;
Carninci P., Hayashizaki Y.;
"High-efficiency full-length cDNA cloning.";
Meth. Enzymol. 303:19-44(1999).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=21085660; PubMed=11217851;
RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RNL Nature 409:685-690(2001).
[3]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RA The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RRL Nature 420:563-573(2002).
[4]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=2049374; PubMed=1107681;
Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RRL Genome Res. 10:1617-1630(2000).
[5]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX MEDLINE=20530913; PubMed=11076861;
Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
Konno H., Akiyama J., Nishi K., Kicsunai T., Tashiro H., Itoh M.,
Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Harada A.,
Fujiwaka S., Inoue K., Togawa K., Izawa M., Ohara E., Watahiki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RRL Genome Res. 10:1757-1771(2000).
[6]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adrenal gland;
RX ADACHI J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
Kato H., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
Saito R., Saion H., Sakai C., Sakai K., Sakazume N., Sano H.,
Saeki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RRL Submitted (APR-2002) to the EMBL/GenBank/DDer databases.
EMBL; AK078659; BAC37350.1; -;
GO; GO:0016021; C:integral to membrane; IEA.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP_1.
Pfam; PF00090; TSP_1; 5.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00209; TSP1; 5.
PROSITE; PS50092; TSP1; 5.
Transmembrane.
NON TER
RP SEQUENCE 478 AA; 51394 MW; F4A9DFFF173F0A376 CRC64;
Query Match 41.2%; Score 507.5; DB 2; Length 478;
Best Local Similarity 31.9%; Pred. No. 1.5e-32;
Matches 91; Conservative 26; Mismatches 81; Indels 87; Gaps 3;

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=NMRI;	
RX	MEDLINE=96414430; PubMed=8817451;	
RA	Adams R.H., Betz H., Puechel A.W.;	
RT	"A novel class of murine semaphorins with homology to thrombospondin	
RT	is differentially expressed during early embryogenesis.";	
RL	Mech. Dev. 57:33-45(1996).	
CC	CC -!- FUNCTION: May act as positive axonal guidance cues.	
CC	CC -!- SUBCELLULAR LOCATION: Type I membrane protein.	
CC	CC -!- TISSUE SPECIFICITY: In adult, only detected in brain.	
CC	CC -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and	
CC	adult tissues. Its abundance decreases from E10 to birth.	
CC	CC -!- SIMILARITY: Belongs to the semaphorin family.	
CC	CC -!- SIMILARITY: Contains 1 Sema domain.	
CC	CC -!- SIMILARITY: Contains 7 TSP type-1 domains.	
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; X97818; Chordata; Craniata; Vertebrata; Euteleostomi;	
DR	HSSP; P07996; ILSL.	
DR	MGI; 107555; Sema5b.	
DR	InterPro; IPR003659; Plexin-like.	
DR	InterPro; IPR002165; Plexin_repeat.	
DR	InterPro; IPR001627; Sema.	
DR	InterPro; IPR000884; TSP1.	
DR	InterPro; IPR008085; TSP_1.	
DR	Pfam; PF01403; Sema; 1.	
DR	Pfam; PF00090; TSP 1; 5.	
DR	PRINTS; PR01705; TSP1REPEAT.	
DR	SMART; SM00423; PSI; 1.	
DR	SMART; SM00630; Sema; 1.	
DR	SMART; SM00209; TSP1; 5.	
DR	PROSITE; PS1004; Sema; 1.	
DR	PROSITE; PS50092; TSP1; 5.	
DR	Developmental protein; Glycoprotein; Multigene family; Neurogenesis;	
KW	Repeat; Signal; Transmembrane.	
FT	SIGNAL 1 19 Potential.	
FT	CHAIN 20 1093 Semaphorin 5B.	
FT	DOMAIN 20 978 Extracellular (Potential).	
FT	TRANSMEM 979 999 Potential.	
FT	DOMAIN 1000 1093 Cytoplasmic (Potential).	
FT	DOMAIN 45 495 Sema.	
FT	DOMAIN 551 605 TSP type-1 1.	
FT	DOMAIN 606 662 TSP type-1 2.	
FT	DOMAIN 664 713 TSP type-1 3.	
FT	DOMAIN 721 776 TSP type-1 4.	
FT	DOMAIN 795 850 TSP type-1 5.	
FT	DOMAIN 852 907 TSP type-1 6.	
FT	DOMAIN 908 952 TSP type-1 7.	
FT	CARBOHYD 59 59 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 95 95 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 157 157 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 178 178 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 287 287 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 333 333 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 378 378 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 532 532 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 539 539 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 547 547 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 602 602 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 728 728 N-linked (GlcNAc. .) (Potential).	
FT	CARBOHYD 944 944 N-linked (GlcNAc. .) (Potential).	

SQ	SEQUENCE 1093 AA; 120326 MW; 295C9B1E8108717 CRC64;	
Query Match	41.2%; Score 507.5; DB 1; Length 1093;	
Best Local Similarity	31.9%; Pred. No. 3.3e-32;	
Matches	91; Conservative 26; Mismatches 81; Indels 87; Gaps 3;	
QY	3 HGFSSWASWRAACVTCGKGIQKSRCLNQLPANGKPCQGSDDLEMRNC-QNKPCPVVG 61	
Db	606 NGAWTANSSWAQSTSCGIGFVQRQSCSNPAPRHGGRICVGRSRERFCNENTPCVPPI 665	
QY	62 SWEWSLWEECTSCGRGNQTRTRTCNNPVSQVHGGRPCENAVEIIMCNTRPCP----- 115	
Db	666 FWASGWSKSCNNCGGVQSRSSCBEN-----GNSCPGCGVEFTCNPEACPEVRNT 719	
QY	116 -----	
Db	720 PWTPLPVNTVQGARQEQPRFTCTRAPLPDPHGLQFGKRTETRTCPADGTGACDADAL 779	
QY	116 -----VHGAWSAWQFWGTCSGCKGTQTRALCNPPPPFGGSGYCDGAET 161	
Db	780 VEDLLRSGSTSPHPLNGWATGWSSCSRDCELGFRVVRKTRCTNPPRNGGLPCVGDAA 839	
QY	162 QMVCNERNCPHKGWATWASWACSVCSCGGARQRTGSDPVP 206	
Db	840 EYDCNFPQACPVRGAWSCWTAWSCQSCGCGHYQRTSCTSPAP 884	
RESULT 14		
QSPCK8	PRELIMINARY; PRT; 1088 AA.	
ID	Q6PCK8	
AC	Q6PCK8;	
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	MGC68835 protein.	
OS	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_TaxID=8355;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Embryo;	
RX	MEDLINE=22388257; PubMed=12477932;	
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,	
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,	
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fatey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,	
RA	Jones S.J., Marra M.A.;	
RA	"Generation and initial analysis of more than 15,000 full-length human	
RT	and mouse cDNA sequences.";	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	
RL	[2]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE=Embryo;	
RC	TISSUE=Embryo;	
RX	MEDLINE=22341132; PubMed=12454917;	
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RT	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative.";	
RL	Dev. Dyn. 225:384-391 (2002).	

```

RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059288; AAH59288.1;
DR InterPro; IPR003659; Pfam-like.
DR InterPro; IPR002165; Pfam-like.
DR InterPro; IPR001627; Pfam-like.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP1.
DR Pfam; PF01437; Pfam; 1.
DR Pfam; PF01403; Pfam; 1.
DR Pfam; PF00090; TSP1; 1.
DR PRINTS; PR01705; TSP1; 1.
DR SMART; SM00423; TSP1; 1.
DR SMART; SM00630; TSP1; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS0092; TSP1; 6.
SQ SEQUENCE 1088 AA; 122657 MW; 4DFCD371A7CD8176 CRC64;

Query Match 41.1%; Score 506.5; DB 2; Length 1088;
Best Local Similarity 31.2%; Pred. No. 3.9e-32;
Matches 89; Conservative 31; Mismatches 78; Indels 87; Gaps 3;

QY 3 HGGFSOWSAWRACSVTCGKGIQKESRLCNOPLPANGKPCQGSDEMRNC-QNKPCPVDG 61
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 SWSWSLWEECTRCGRGNQTRTTCNNPSPVQHGRCPCGNNAVEIIMCNIIRPCP- 115
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 662 FWTSGSWTKCSDCGGTHSRQSCEN-----GNTCPGCTMEFRTCNPEPCPEVRNT 715
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 ----- 115
DB 716 PWTWMPVNTTQGARQEQFRYTCRAQLADPHNLQFRRKKTETTRPCPDNGSAACEMDTL 775
QY 116 -----VHGAWSAWQPGWTCSECGKGTQTRARLNNPPPPAFGGSYCDGAET 161
DB 776 VDDLNRNKTSGHIINGWSLWSWSSCSRDCQGFPRSRKRTCSNPEPRKGGLPCTGSAM 835
QY 162 QMQVCNERNCPIHGKWATWASWSACSVCSSGGRARQTRGCDPVP 206
DB 836 EYQDCNPQPCPVKGSWCSWSPWTCQCATCGGGHYQRTCTCTNPPP 880

RESULT 15
AAH59288 PRELIMINARY; PRT; 1088 AA.
AC AAH59288;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC68835 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

```

```

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC059288; AAH59288.1;
SQ SEQUENCE 1088 AA; 122657 MW; 4DFCD371A7CD8176 CRC64;

Query Match 41.1%; Score 506.5; DB 2; Length 1088;
Best Local Similarity 31.2%; Pred. No. 3.9e-32;
Matches 89; Conservative 31; Mismatches 78; Indels 87; Gaps 3;

QY 3 HGGFSOWSAWRACSVTCGKGIQKESRLCNOPLPANGKPCQGSDEMRNC-QNKPCPVDG 61
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 SWSWSLWEECTRCGRGNQTRTTCNNPSPVQHGRCPCGNNAVEIIMCNIIRPCP- 115
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 662 FWTSGSWTKCSDCGGTHSRQSCEN-----GNTCPGCTMEFRTCNPEPCPEVRNT 715
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 ----- 115
DB 716 PWTWMPVNTTQGARQEQFRYTCRAQLADPHNLQFRRKKTETTRPCPDNGSAACEMDTL 775
QY 116 -----VHGAWSAWQPGWTCSECGKGTQTRARLNNPPPPAFGGSYCDGAET 161
DB 776 VDDLNRNKTSGHIINGWSLWSWSSCSRDCQGFPRSRKRTCSNPEPRKGGLPCTGSAM 835
QY 162 QMQVCNERNCPIHGKWATWASWSACSVCSSGGRARQTRGCDPVP 206
DB 836 EYQDCNPQPCPVKGSWCSWSPWTCQCATCGGGHYQRTCTCTNPPP 880

Search completed: November 17, 2004, 15:55:32
Job time : 197 secs

```